Measurement of the mass of the Higgs boson in the $H \to ZZ^* \to 4\ell$ channel at the ATLAS detector

The University Of Sheffield.

Thomas Powell

Department of Physics and Astronomy
University of Sheffield

This thesis is submitted for the degree of
Doctor of Philosophy

June 2020

Abstract

Presented in this thesis are the results from a measurement of the mass of the Higgs Boson in the $H \to ZZ^* \to 4\ell$ ($\ell = e, \mu$) channel using ATLAS data from pp collisions at a centre of mass energy of 13 TeV corresponding to an integrated luminosity of 139 fb^{-1}. This measurement uses a Double-sided Crystal Ball function as an analytic signal model, incorporating per-event uncertainties as estimated using a quantile regression neural network. The mass of the Higgs boson is determined to be 124.92 $\pm$ 0.19 (stat.) $^{+0.09}_{-0.06}$ (syst.) GeV, in line with the previous LHC Run 1 combined result from the ATLAS and CMS collaborations of 125.09 $\pm$ 0.21 (stat.) $\pm$ 0.11 (syst.) GeV and the previous ATLAS measurement, which used 36 fb^{-1} of 13 TeV pp collision data, of 124.97 $\pm$ 0.18 (stat.) $\pm$ 0.20 (syst.) GeV.

Also presented are results from a method to validate the ATLAS photon energy calibration procedure by calibrating electrons from $Z \to ee$ events as photons. This technique is also used to calibrate unconverted photons as converted and vice versa to investigate the effect of misidentifying photon conversions. From this it is found that misidentifying all unconverted photons as converted photons which leave either one or two tracks in the transition radiation tracker (single/double TRT converted photons) leads to an η-independent bias in calibrated energy of approximately 1%. It is also found that misidentifying all single/double TRT converted photons as unconverted leads to a smaller bias, which is around 1% in the pseudorapidity region $0.8 \leq |\eta| < 1.37$.

Results from the application of shift and width adjustments to simulated electron shower shapes to correct for mismodelling are presented and a new method for performing such corrections using a Gaussian convolution technique is demonstrated. This new method is found to reduce the amount of over-correction in the application of corrections to electron shower shapes.

Acknowledgements

First of all, a massive thanks to my supervisor, Christos, for his supervision and support during my time working at Sheffield.

Next, to the members of the Sheffield ATLAS group for the friendly and welcoming atmosphere.

I would also like to acknowledge the inhabitants of D44 for keeping me sane over the years and providing an environment that is (usually) conducive to work.

Thanks also to those who went with me on adventures to the most amazing mountainous (or otherwise) places during my time at CERN. It certainly helped give me a much needed distraction!

In a similar light, thank you to all my climbing friends in Sheffield. I don't think there are many better ways to keep both the physical and mental health in check!

Thanks to my flatmate Rory for his support (and cooking!) during these past few months and to Dan for the long and fascinating Monday evening discussions with us (not to mention baking excellent flapjacks!).

Finally to my parents, for their love and support throughout it all.

"There is a theory which states that if ever anyone discovers exactly what the Universe is for and why it is here, it will instantly disappear and be replaced by something even more bizarre and inexplicable. There is another theory which states that this has already happened."

Douglas Adams

Table of contents

List of figures **xi**

List of tables **xix**

1 Introduction **1**

2 Theory **5**

2.1 Particle content of the Standard Model 5
2.2 Quantum electrodynamics (QED) 7
2.3 Quantum chromodynamics (QCD) 8
2.4 The weak interaction 9
2.5 Spontaneous symmetry breaking and the Higgs mechanism 10
2.6 Electroweak symmetry breaking in the Standard Model 13
2.7 Production of Higgs bosons in hadronic collisions 16
2.8 Four-lepton decays of Higgs bosons 18

3 The Large Hadron Collider and the ATLAS experiment **21**

3.1 The Large Hadron Collider 21
3.2 The ATLAS detector 21
3.2.1 Inner detector 22
3.2.2 Calorimetry 24
3.2.3 Muon system 30
3.2.4 Trigger system 30

4 Electron and photon reconstruction and calibration **33**

4.1 Electron and photon reconstruction 33
4.1.1 EM cluster reconstruction 34
4.1.2 Track reconstruction 35
4.1.3 Track-cluster matching 35

4.1.4 Supercluster formation . . . 36
4.2 Electron-photon calibration . . . 36
4.2.1 Overview of calibration procedure . . . 36
4.2.2 Further checks on the photon calibration procedure . . . 41
4.2.3 Investigating the effect of misidentifying photon conversions . . . 45
4.3 Electron identification . . . 47
4.3.1 Likelihood discriminant . . . 47
4.3.2 Shower shape corrections . . . 50
4.3.3 Improvements to shower-shape corrections . . . 54

5 Standard Model Higgs boson mass measurement in $H \to 4\ell$ at 36.1 fb^{-1} 57
5.1 Overview . . . 57
5.2 Simulation of events . . . 58
5.3 Object requirements and event selection . . . 59
5.3.1 Muons . . . 59
5.3.2 Electrons . . . 60
5.3.3 Event selection . . . 60
5.4 Improvements to $m_{4\ell}$ resolution . . . 61
5.5 Category definitions . . . 62
5.6 Background model . . . 62
5.7 Signal model . . . 64
5.7.1 Per-event response method . . . 64
5.7.2 Template method . . . 66
5.8 Results . . . 67

6 Standard model Higgs Boson mass measurement in $H \to 4\ell$ at 139 fb^{-1} 71
6.1 Analysis strategy . . . 71
6.2 Signal model . . . 72
6.2.1 Analytic parameterisation . . . 72
6.2.2 Conditional parameterisation using per-event errors . . . 92
6.3 Background model . . . 95
6.3.1 Background modelling . . . 95
6.4 Signal and background normalisation . . . 95
6.5 Systematic uncertainties . . . 99

7 Results of the Higgs boson mass measurement at 139 fb^{-1} 107
7.1 Expected results . . . 107

7.1.1 PER model 107
7.1.2 Model without per-event resolution 108
7.1.3 Compatibility between models 110
7.2 Observed results 110
7.2.1 PER model 112
7.2.2 Model without per-event resolution 115

8 Conclusion **117**

References **119**

List of figures

2.1 The potential of Equation 2.23 for the cases $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right). 11
2.2 The dominant processes for Higgs boson production in proton-proton collisions. 17
2.3 Higgs boson production cross section vs m_{H} for pp collisions with a centre of mass energy of 13 TeV. [15] 18
2.4 SM Higgs boson branching ratio of various decay modes vs m_{H}. [15] . . 19
2.5 The Feynman diagram for the Higgs decay to four leptons ($H \to ZZ^* \to 4\ell$) (for $l = e, \mu$). The "*" indicates the off-shell Z. 20

3.1 Cutaway diagram of the ATLAS detector. [17] 22
3.2 Diagram showing the elements of ATLAS tracking system. [18] 23
3.3 Schematic diagram of an electromagnetic shower initiated by an electron. The incident electron radiates a Bremsstrahlung photon, which itself creates an electron positron pair through the process of pair production. Note that each of these vertices is initiated by the exchange of a photon with an atom of the detector material, which is not shown. 24
3.4 The three dominant processes by which electrons and photons lose energy at medium to low energies ($\lessapprox 10$ MeV). 25

3.5 The cross sections for each of the processes that contribute to photon scattering vs incident photon energy. The dominant process at low energies is the photoelectric effect, the cross section for which is denoted τ here. The Compton effect dominates at energies of $\sim$ 1 MeV, the cross section for which is denoted σ_{INCOH} . Above these energies, pair production from interactions with the atomic nuclei dominates, κ_{n}. The cross sections for sub-leading processes of Rayleigh scattering (σ_{COH}), pair production from atomic electron interaction (κ_{e}) and nuclear photoabsorption ($\sigma_{\mathrm{PH.N}}$) are also shown along with the total cross section from experiment (σ_{TOT}). [20] 26

3.6 The longitudinal difference between electron initiated showers in different materials (a) and photon and electron initiated showers in tin (b). The fraction of initial energy deposited per X_0 is shown for a 10 GeV electron in lead, iron and aluminium absorbers (a) and per 0.5 X_0 for 10 GeV photons and electrons in a tin absorber (b). These were both calculated from Monte Carlo simulations. [19, 21] 27

3.7 A diagram of the layout of the ATLAS calorimeters. [17] 29

3.8 A diagram of one of the barrel modules of the ATLAS electromagnetic calorimeter. [17] 29

3.9 A cutaway diagram of the ATLAS muon system. [17] 30

4.1 Flowchart showing an overview of the electron and photon calibration procedure. [32] 37

4.2 $(\langle E_e/E_{\mathrm{raw}}\rangle^{\mathrm{MC}} - \langle E_e/E_{\mathrm{raw}}\rangle^{\mathrm{data}}) - (\langle E_{\gamma^{\mathrm{TR}}}/E_{\mathrm{raw}}\rangle^{\mathrm{MC}} - \langle E_{\gamma^{\mathrm{TR}}}/E_{\mathrm{raw}}\rangle^{\mathrm{data}})$ vs (a) fraction recovered energy in the EM calorimeter deposited in the pre-sampler, E_0/E_{tot}, (b) ratio of energy deposited in the first layer of the EM calorimeter to the second layer, E_1/E_2, (c) $|\eta|$ and (d) ratio of the energy deposited in the 3×3 cells centred at the cluster to the deposit in 3×7, R_ϕ. The choice of y-axis scale here is to allow the same scale to be used as in Figure 4.4. 42

4.3 Distributions of kinematic and shower shape variables for electrons from $Z \to ee$ MC and photons from single photon MC. All figures show $Z \to ee$ with selections of single lepton triggers, exactly two electrons satisfying medium identification requirements, with both having $E_{\mathrm{T}} > 25$ GeV and exactly one topocluster and photon selections of a single unconverted photon with $E_{\mathrm{T}} > 25$ GeV satisfying strict identification criteria. Figures (b) and (c) have the additional requirement of $E_{\mathrm{T}} < 50$ GeV for both electrons and photons. In each figure, the filled region is the distribution for electrons and the empty circles are that of the photons. 43

4.4 The difference the mean ratio of calibrated to raw energy between transformed and real photons, $\langle E_{\gamma^{\mathrm{TR}}}/E_{\mathrm{raw}}\rangle^{Z\to ee} - \langle E_{\gamma}/E_{\mathrm{raw}}\rangle^{\mathrm{single}-\gamma}$, vs shower shape variables. The variation for each variable is shown in four bins of $|\eta|$. 44

4.5 Behaviour of $E_{\gamma}^{\mathrm{transformed}}/E_{\mathrm{raw}} - E_{\gamma}^{\mathrm{original}}/E_{\mathrm{raw}}$, where E_{raw} is the uncalibrated (raw) energy of the photon, E_{original} refers to the calibrated energy of MC unconverted photons and $E_{\gamma}^{\mathrm{transformed}}$ to that of the corresponding photon transformed to either a single or double TRT converted photon with 50% probability. (a) Shows the distribution and (b) shows the behaviour vs $|\eta|$. 45

4.6 Behaviour of $E_{\gamma}^{\mathrm{transformed}}/E_{\mathrm{raw}} - E_{\gamma}^{\mathrm{original}}/E_{\mathrm{raw}}$, where E_{raw} is the uncalibrated (raw) energy of the photon, E_{original} refers to the calibrated energy of the MC single/double TRT converted photon and $E_{\gamma}^{\mathrm{transformed}}$ to that of the corresponding photon transformed to an unconverted photon. (a) Shows the distribution and (b) shows the behaviour vs $|\eta|$. 46

4.7 Corrections applied to the variable R_{had} for $4 \leq E_{\mathrm{T}} < 7$ GeV the bin (a) $0.00 \leq |\eta| < 0.60$, (b) $0.60 \leq |\eta| < 0.80$ and (c) $0.80 \leq |\eta| < 1.15$. The data and MC used correspond to $J/\psi \to ee$ decay events. This variable requires a relatively large width correction and a small shift. 50

4.8 Corrections applied to the variable f_3 for $7 \leq E_{\mathrm{T}} < 10$ GeV left to right in the bin (a) $0.80 \leq |\eta| < 1.15$, (b) $1.52 \leq |\eta| < 1.81$ and (c) $1.81 \leq |\eta| < 2.01$. The data and MC used correspond to $J/\psi \to ee$ decay events. This variable requires a minor shift and width correction. 51

4.9 Corrections applied to the variable f_3 for $30 \leq E_{\mathrm{T}} < 40$ GeV in the bin (a) $0.00 \leq |\eta| < 0.60$, (b) $0.60 \leq |\eta| < 0.80$ and (c) $1.81 \leq |\eta| < 2.01$. The data and MC used are from $Z \to ee$ decays. This variable requires a minor shift and width correction. 52

4.10 Variation in corrections vs $|\eta|$ and E_{T} for f_3 from $J/\psi \to ee$ decays. Shown in (a) and (b) are the variation in shift and width correction vs E_{T} for $0.60 \leq |\eta| < 0.80$ respectively and (c) and (d) shift and width correction vs $|\eta|$ for $7 \leq E_{\mathrm{T}} < 10$ GeV respectively. 53
4.11 Electron ID efficiency ($\varepsilon_{\mathrm{id}}$) vs E_{T} (left) and η (right) for the Loose, Medium and Tight operating points [27]. The comparison with expected efficiencies from MC are also shown. 54
4.12 Comparison of the two correction methods shown for the distribution of f_3 in the $15 \leq E_{\mathrm{T}} < 20$, $0.60 \leq |\eta| < 0.80$ bin. These are the shift and FWHM method (left) and the convolution method (right). 56
4.13 Comparison of the two correction methods shown for the distribution of R_{had} in the $15 \leq E_{\mathrm{T}} < 20$, $0.60 \leq |\eta| < 0.80$ bin. These are the shift and FWHM method (left) and the convolution method (right). 56

5.1 Examples of KDE smoothing for the case of fixed-size kernels (a) and adaptive (b). 63
5.2 A visualisation of the Gaussian mixture reduction procedure. [41] . . . 65
5.3 Example of the morphing of template $m_{4\ell}$ distributions in the 4μ channel. [41] 67
5.4 Results for the Higgs boson mass measurement at 36 fb^{-1} in the $H \to 4\ell$ channel. Left, the projection of the fit from all 16 categories and right, the profile likelihood of the fit to each channel. [6] 68
5.5 Combination of ATLAS Run 1 and early Run 2 measurements of m_{H} in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels. This is compared to the combined ATLAS + CMS result from Run 1, with which good agreement is seen. [6] 69

6.1 Fits of DCB (left column) and CBG (right column) models to Monte Carlo simulations for $m_{\mathrm{H}} = 125$ GeV, ggF production mode in each channel shown in a linear scale. The DCB model fits better in all channels. 75
6.2 Fits of DCB (left column) and CBG (right column) models to Monte Carlo simulations for $m_{\mathrm{H}} = 125$ GeV, ggF production mode in each channel shown in a log scale. The DCB model fits better in all channels. 76
6.3 Left: α_{high} vs x_{max} and right: n_{high} vs x_{max} for each channel. The α_{high} and n_{high} parameters tend to reach more stable values with increasing x_{max}. 78

6.4 Left: α_{low} vs x_{low} and right: n_{low} vs x_{min} for each channel. The α_{low} and n_{low} parameters tends to reach more stable values with decreasing x_{min}. 79
6.5 Validation of the DCB parameterisation on each channel, shown in a linear and a log scale. The DCB is fitted to the second set with μ floating and all other parameters fixed. From these fits, Equation 6.7 is used to estimate m_{H}. 81
6.6 The measured values of m_{H} from validation fits in each of the four channels as well as the combined fit across all channels. The uncertainties here reflect the expected statistics at 139 fb^{-1}. 82
6.7 Fitted μ vs true m_{H} from fitting a DCB model to MC mass points in the range 123 to 127 GeV for each channel. The μ parameter shows a strong linear relationship with m_{H}; however, the slope is not exactly 1 in each case. 82
6.8 All DCB parameters (except μ) vs m_{H} in the 4μ channel. The α_{high} and n_{high} parameters show some dependence on m_{H} whereas all other parameters show no strong dependence. 83
6.9 Bias in m_{H} measurement in 4μ channel from varying each parameter of the model by $\pm 1\sigma$, where σ is its uncertainty. Starting from the nominal model, one parameter is increased by $+1\sigma$ from its nominal value, the model is fitted to the $m_{\text{H}}^{\text{true}} = 125$ GeV MC set and the value of $m_{\text{H}}^{\text{measured}}$ is plotted. This process is repeated for the $+1\sigma$ case and for each parameter in the model. The value of $m_{\text{H}}^{\text{measured}}$ from the nominal model is shown as well as its uncertainty. 84
6.10 Optimisation of the BDT binning, showing the values of bin significance, z vs number of bins for the even binning (blue) and optimised binning (red) cases. 85
6.11 Distributions of $m_{4\ell}$ in each of the four BDT_{ZZ} bins in each channel for signal ggF. The mean of the distribution tends to be closer to m_{H} and the RMS lower for the more signal-like bins. 86
6.12 Calibration plots for the analytic model in the 4μ channel, BDT bin 4. 87
6.13 Values of $m_{\text{H}}^{\text{measured}}$ from each category as well as one from a combined fit across all 16 categories. The fits show good closure. Uncertainty bands in this plot reflect the expected statistics, rather than the number of generated MC events. 88

6.14 Pull distributions for $m_{\mathrm{H}}^{\mathrm{pull}}$ from fitting the model for $m_{\mathrm{H}}^{\mathrm{true}} = 125$ GeV to (a) bootstrapped MC and (b) toy data generated from the model. In both cases, all parameters other than m_{H} are fixed. In each case a Gaussian is fitted to the distribution. The distribution from toys shows no bias and good calibration of uncertainties and the bootstrapped shows a small negative bias and some underestimation of uncertainties. 90

6.15 Results for bootstrapped MC using the ggF, VBF and VH production modes. Shown are (a) Distribution of $m_{\mathrm{H}}^{\mathrm{measured}}$ and (b) distribution of the uncertainties in $m_{\mathrm{H}}^{\mathrm{measured}}$ from fitting the model for $m_{\mathrm{H}}^{\mathrm{true}} = 125$ GeV 10,000 times to bootstrapped MC. All parameters other than m_{H} are fixed. A Gaussian fit is added to (a) with an expected mean of 125 GeV. The plot shows a negative bias in line with Figure 6.14. (b) shows that the most probable uncertainty in $m_{\mathrm{H}}^{\mathrm{measured}}$ is 153 MeV with the mean uncertainty being 155 MeV. 90

6.16 Distributions of $m_{\mathrm{H}}^{\mathrm{measured}}$ from validating the models for $m_{\mathrm{H}}^{\mathrm{true}} = 124$, 125.5 and 126 GeV on bootstrapped MC for their respective mass points. Each show a bias of up to 5 MeV. 91

6.17 Calibration plots for the PER model in the 4μ channel, BDT bin 4. . . 94

6.18 Demonstration of the effect of varying the smoothing parameter, ρ, on the $m_{4\ell}$ distribution in the 4μ channel for the qqZZ background, shown for values of $\rho = 1, 2, 3.4$. In the case of $\rho = 1$ (a), nearly all features are preserved but the shape is susceptible to fluctuations. In the case of $\rho = 3.4$ (c), these fluctuations are smoothed over, but some of the shape is lost in doing so. Note that the sharp drop around $m_{4\ell} \approx 150$ GeV is not physical, and is caused by the filtering of events. 97

6.19 Comparison of different techniques to smooth the two MC samples for $qqZZ$ in the 4μ channel. The first, filtered, sample covers the range 100-150 providing greater MC statistics here. The second, the inclusive sample, covers all other parts of the $m_{4\ell}$ spectrum. The first approach is to smooth each component individually (red) and the second is to smooth the components together (blue). 98

6.20 Systematic variations of $m_{4\ell}$ distributions for the signal model of the electron energy scale systematic in each channel and inclusive of all channels. 101

6.21 Systematic variations of $m_{4\ell}$ distributions for the signal model of the muon momentum scale systematic in each channel and inclusive of all channels. 102
6.22 Pull of μ from fitting a DCB to a toy dataset generated from that same DCB with μ floating and all other parameters fixed. A Gaussian fit to the distribution is also shown. The parameters of this Gaussian show a bias in the negative direction. 103
6.23 Mean of the fitted Gaussian to the distribution of pulls on μ from fitting a DCB to toy data with μ floating, against number of events. The bias is expected to be proportional to $1/n_{evts}$, so a function of the form $a + b/n_{evts}$ is fitted, with the expectation that $a = 0$, where both a and b are measured in GeV. It is found that $a < 0$, so the bias is not proportional to $1/n_{evts}$. 104
6.24 Bias from pulls of μ of a CB fit to toy data generated from the same CB, against number of events. The different sets of points correspond to different parameters of the CB floating/fixed. The closed circles are for the case of only μ floating (with all parameters fixed), open triangles for μ and σ floating and open circles for μ, σ and α floating. A fit of the form $a \times n_{evts}^{b}$ is also shown. 105
6.25 Distribution of m_{H} values from fitting the calibrated analytic model to toy data generated from the model. This shows that the bias due to the estimator is 0 within uncertainty. 106

7.1 Expected results for the PER model (a) results from fits to the full MC statistics and (b) distribution of m_{H} from fitting to bootstrapped MC with a Gaussian fit. 108
7.2 Expected systematic ranking for the PER model. The red and blue bars indicate the prefit and postfit effects of a particular systematic on the uncertainty respectively (top x-axis). The black points indicate the amount by which the corresponding nuisance parameter is pulled from its nominal value (bottom x-axis). 109
7.3 Expected results for the model without accounting for per-event resolution (a) results from fits to the full MC statistics and (b) distribution of m_{H} from fitting to bootstrapped MC with a Gaussian fit. 109
7.4 Expected systematic ranking for the model without accounting for per-event resolution. 110

7.5 Compatibility between the model with and without per-event resolution. (a) shows the difference in measured m_{H} between when fitting to the same sets of bootstrapped MC. As the peak is centred on 0, no bias is observed with a standard deviation of 45 MeV. (b) shows the difference in uncertainty between the same sets of bootstrapped MC. On average, the accounting for per-event resolution improves uncertainty estimation by 4 MeV with a standard deviation of 5 MeV. 111
7.6 Distributions of the $m_{4\ell}$ spectrum (left) and σ_i (right). The filled regions represent the expected distributions from MC simulation and the black points represent the observed distributions from data. In both cases, the MC and data are in agreement. [46] 112
7.7 Projection of the PER fitted across all 16 categories. The black points represent the ATLAS 139 fb^{-1} data, the red line, the fit, and the shaded region, the background model. [46] 113
7.8 Results from the PER model, per channel and combined. Shown are the measured values of m_{H} (left) and the likelihood scans (right) for each channel and combined. Each of the four per-channel fits is performed simultaneously across four BDT bins and the combined fit is performed simultaneously across all 16 categories. [46] 113
7.9 Observed systematic ranking for the PER model. 114
7.10 Expected distribution of uncertainties on m_{H} when accounting for (black) and when not accounting for per-event resolution (blue). The solid lines indicate the distribution of uncertainties from the fits to bootstrapped MC sets and the dashed lines, the observed values of the uncertainties. [46] 116

List of tables

2.1 A list of all of the standard model bosons [13]. 6

2.2 A list of all the Standard model fermions. All fermions have an intrinsic spin of 1/2 and are differentiated from one another by their charges, masses and their interactions with the three forces [13, 14]. Note that the values for the neutrino masses are direct limits, stronger limits are achieved from indirect measurements neutrino oscillation measurements. Also note that in the case of the neutrinos, the mass eigenstates are different from the flavour eigenstates. 6

4.1 Variables used in the training of the likelihood discriminant. Variables defined in e.g. "3×3 cells" refer to the cells of the second layer of the electromagnetic calorimeter in $\Delta\eta \times \Delta\phi$ space in steps of 0.025. The two rightmost columns indicate whether shift or width corrections are applied to the PDF of the variables to correct for shower shape mis-modelling. 48

4.2 Bin boundaries in E_{T} used for electron identification. For the lower E_{T} bins ($E_{\mathrm{T}} \leq 15$ GeV), $J/\psi \to ee$ events are used and $Z \to ee$ for the higher E_{T} (> 15 GeV). 49

4.3 Bin boundaries in $|\eta|$ used for electron identification. 49

5.1 Summary of the event selection for the measurement of m_{H} at 36 fb^{-1}. [38] 60

5.2 Expected and observed yields for events within $115 < m_{4\ell} < 130$ GeV using ATLAS early Run 2 data. The expected yields for signal are calculated assuming $m_{\mathrm{H}} = 125$. [38] 61

6.1 A list of all the simulated signal processes. MC simulations for the two leading processes, ggF and VBF are simulated at several different mass points as these are used for building the model. The sub-leading processes are only used in validation of the model and determining the expected event yield. N_{MC} is the number of unweighted MC events which pass selection and N_{expt} is the expected number of events at 139 fb^{-1} in the $105 < m_{4\ell} < 160$ GeV range. 73

6.2 Summary of results of Figures 6.15 and 6.16. Means, uncertainties and biases of the distributions of $m_{\text{H}}^{\text{measured}}$ from fitting the model for various different values of $m_{\text{H}}^{\text{true}}$ 10,000 times. 92

6.3 A list of all the simulated background processes. For the backgrounds $qqZZ$, $ggZZ$, $\text{t}\bar{\text{t}}\mu\mu$ and $\text{t}\bar{\text{t}}ee$, multiple samples are used with a filter on $m_{4\ell}$ applied. This is so MC samples with enhanced statistics in a desired range can be generated. Overlap between such samples is accounted for by the weighting of events. N_{MC} is the number of unweighted MC events which pass selection and N_{expt} is the expected number of events at 139 fb^{-1} in the $105 < m_{4\ell} < 160$ GeV range. 96

7.1 Expected and observed yields for events within $115 < m_{4\ell} < 130$ GeV. The expected yields for signal are calculated assuming $m_{\text{H}} = 125$ GeV. [61]111

7.2 Normalisations for signal and background for the PER model. A small excess of background and small under-fluctuation of signal are observed. 114

Chapter 1

Introduction

Electroweak symmetry breaking and the Higgs mechanism were proposed in 1964 as a method by which the W and Z bosons of the Standard Model (SM) could acquire mass [1, 2, 3]. In doing so, this predicted the existence of a boson which came to be known as the *Higgs Boson*. The Large Hadron Collider (LHC), currently the largest particle collider in operation, was built in order to search for this boson by colliding protons at energies up to 14 TeV. In 2012, the ATLAS and CMS experiments, using data from LHC collisions, successfully verified the existence of the Higgs Boson [4, 5]. Following the verification of its existence, work has since moved to measuring its properties to determine whether it is exactly the boson predicted by the Standard Model.

Presented here are studies to determine one of these properties, the mass, m_{H}, using data from the ATLAS experiment. Whilst the mass is a free parameter in the SM, it is nonetheless an important parameter to measure. The reason for this is that other parameters such as the branching ratio of various decay modes depend on the mass, a dependence which is specified by the SM. So, by measuring the mass along with other parameters, the SM can be tested.

Measuring m_{H} requires the decay products to be reconstructed accurately. For this reason, the decay of a Higgs Boson to a pair of Z bosons which themselves decay to either electrons or muons (denoted $H \to ZZ^* \to 4\ell$), is particularly suited to such a measurement. Whilst the branching ratio to this decay mode is considerably lower than that of dominant decay modes, this decay is particularly suited for two reasons. Firstly, the decay products (electrons and muons) are precisely reconstructed by the ATLAS detector. Secondly, the signal is a sharp peak atop a smooth background, allowing precise measurement of the Higgs Boson properties.

The previous measurement of the mass of the Higgs boson by ATLAS in the $H \to ZZ^* \to 4\ell$ channel used proton-proton collision data collected during the LHC

early Run 2 data taking period in 2015 and 2016, corresponding to an integrated luminosity of 36.1 fb^{-1} and a centre of mass energy of 13 TeV. The mass of the Higgs boson was determined to be 124.79 $\pm$ 0.37 GeV [6]. This measurement was obtained using a *per-event response* signal model which describes the individual lepton energy and momentum responses by the weighted sum of Gaussians (described in Section 5.7). A measurement using ATLAS early Run 2 data was also performed in the $H \to \gamma\gamma$ channel, for which an analytic signal model was used. The result from this measurement was 124.93 $\pm$ 0.40 GeV. The current most precise measurement of the mass of the Higgs boson from ATLAS, using results from both the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels from full Run 1 and early Run 2 data is 124.97 $\pm$ 0.24 GeV [6]. CMS also measured the mass in the $H \to ZZ^* \to 4\ell$ channel using data collected during early Run 2, corresponding to an integrated luminosity of 35.9 fb^{-1}. This returned a value of 125.26 ± 0.21 GeV [7], in agreement with those determined by ATLAS. The current most precise measurement of the mass comes from the combination of the results from CMS using data collected during LHC Run 1 and early Run 2 from both the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels. The result of this is $m_{\mathrm{H}} = 125.38 \pm 0.14$ GeV [8].

For the measurement of m_{H} using the full ATLAS LHC Run 2 dataset, a new signal model is used, the results from which are presented here. The ATLAS full Run 2 dataset contains data from 13 TeV pp collisions corresponding to an integrated luminosity of 139 fb^{-1} collected between 2015-2018. This new method aims to describe the four-lepton invariant mass distribution using an analytic function. This method is less computationally intensive than the per-event response approach and allows additional parameters of interest to be added more easily than in the case of a template method.

Chapter 2 provides an overview of the theoretical background for the work presented here, describing each of the fundamental forces and showing the need for electroweak symmetry breaking and the Higgs mechanism and how this predicts the existence of the Higgs boson.

Chapter 3 describes the experimental apparatus used to collect the data for all the results presented. Each of the subsystems of the ATLAS detector are described and a brief overview of the ATLAS trigger system is provided.

Chapter 4 describes how electrons and photons are reconstructed and calibrated in the detector. Shown also are results from a study by the author to verify the procedure for the calibration of photon energy and an investigation into the effect of misidentifying photon conversions. The procedure to identify electrons is also described, as well as results from the application of corrections to account for the mismodelling

of electron shower shapes in the ATLAS calorimeters as performed by the author. Preliminary results from the development by the author of a new method to determine such corrections using a method of Gaussian convolution are also presented.

Chapter 5 gives an overview of the ATLAS measurement of the mass of the Higgs boson in the $H \to ZZ^* \to 4\ell$ channel using ATLAS data from early Run 2 corresponding to an integrated luminosity of 36 fb^{-1} from pp collisions at a centre of mass energy of 13 TeV.

Chapter 6 describes the most recent measurement of the Higgs boson mass in the $H \to ZZ^* \to 4\ell$ channel using the full ATLAS Run 2 dataset corresponding to an integrated luminosity of 139 fb^{-1} from pp collisions at a centre of mass energy of 13 TeV to which the author made significant contributions. For this measurement, an analytic description of the signal lineshape is used in the form of a Double-sided Crystal Ball function. Also incorporated into this is a per-event resolution, calculated using machine learning techniques.

Chapter 7 details the results from this mass measurement, providing both expected and observed results and Chapter 8 summarises all the work presented.

Chapter 2

Theory

This chapter gives a brief overview of the Standard Model of particle physics (SM) and draws particular attention to the Higgs mechanism and its predictions. The content here is based upon references [9], [10], [11] and [12].

The SM is a quantum field theory which describes all fundamental particles and their interactions via the strong, weak and electromagnetic forces. So far, it has been very successful in its predictions which have been verified experimentally. 2012 saw the experimental verification of the Higgs Boson, which was until then, the only fundamental particle predicted by the SM that had not been observed [4, 5].

Each of the three fundamental interactions described by the SM will be introduced individually, as well as the unification of the electromagnetic and weak forces into a single electroweak force and the Higgs mechanism.

2.1 Particle content of the Standard Model

The particles of the standard model can be classified into two distinct classes depending on their intrinsic angular momentum or *spin*. These are *bosons* which, in units of $\hbar$, have integer spin and *fermions* with half-integer spin. All Standard Model bosons are listed in Table 2.1.

The fermions are subdivided into quarks and leptons. The quarks are the only fermions that interact via the strong force and are further subdivided into the "up type" and "down type", which have charges of $+\frac{2}{3}e$ and $-\frac{1}{3}e$ respectively. As well as interacting via the strong force, all quarks additionally interact via both the weak and electromagnetic forces. The leptons are subdivided into the charged leptons and the neutrinos. The charged leptons interact both electromagnetically and weakly, whereas the neutrinos only interact weakly. Each of these lowest subdivisions of fermions

Name	Symbol	Spin ($\hbar$)	Charge (e)	Mass GeV
Photon	γ	1	0	0
Gluon	g	1	0	0
W boson	$W^{\pm}$	1	± 1	80.379 ± 0.012
Z boson	Z	1	0	91.1876 ± 0.0021
Higgs boson	H	0	0	125.10 ± 0.14

Table 2.1 A list of all of the standard model bosons [13].

Quarks					
Type	Name	Symbol	Spin ($\hbar$)	Charge (e)	Mass MeV
Up-type	Up quark	u	1/2	+2/3	$2.16^{+0.49}_{-0.26}$
	Charm quark	c	1/2	+2/3	$(1.270 \pm 0.20) \times 10^3$
	Top quark	t	1/2	+2/3	$(172.9 \pm 0.4) \times 10^3$
Down-type	Down quark	d	1/2	−1/3	$4.67^{+0.48}_{-0.17}$
	Strange quark	s	1/2	−1/3	93^{+11}_{-5}
	Bottom quark	b	1/2	−1/3	$(4.18^{+0.03}_{-0.02}) \times 10^3$
Leptons					
Type	Name	Symbol	Spin ($\hbar$)	Charge (e)	Mass MeV
Charged leptons	Electron	e	1/2	−1	0.511
	Muon	μ	1/2	−1	105.66
	Tau	τ	1/2	−1	1776.86 ± 0.12
Neutrinos	Electron neutrino	ν_e	1/2	0	$< 1.1 \times 10^{-6}$
	Muon neutrino	ν_μ	1/2	0	< 0.19
	Tau neutrino	ν_τ	1/2	0	< 18.2

Table 2.2 A list of all the Standard model fermions. All fermions have an intrinsic spin of 1/2 and are differentiated from one another by their charges, masses and their interactions with the three forces [13, 14]. Note that the values for the neutrino masses are direct limits, stronger limits are achieved from indirect measurements neutrino oscillation measurements. Also note that in the case of the neutrinos, the mass eigenstates are different from the flavour eigenstates.

contains three particles, which differ only by their mass. These are referred to as the *generations*, the first generation referring to the lightest fermion in the subdivision and the third generation referring to the heaviest. All the standard model fermions are summarised in Table 2.2.

2.2 Quantum electrodynamics (QED)

Quantum electrodynamics (QED) is a quantum field theory, formed from the Abelian $U(1)_Q$ group, that provides a description of the electromagnetic interaction. This interaction affects all charged particles, i.e. all quarks, the charged leptons e, μ and τ as well as the $W^{\pm}$ bosons. The dynamics of particles and their corresponding fields will be described in the context of Lagrangians. The Lagrangian of a discrete system is given by

$$L = T - V, \tag{2.1}$$

where T and V are the kinetic and potential energies of all the particles in the system. When describing the dynamics of fields, it is necessary to use Lagrangian densities, $\mathcal{L}$, which are related to the discrete form by

$$L = \int \mathcal{L} d^3\mathbf{x}, \tag{2.2}$$

where $d^3\mathbf{x}$ is a volume element. For convenience, Lagrangian densities from here on will be referred to simply as Lagrangians.

The motion of a free, spin-half particle in the absence of an external field is described by the Dirac equation, for which the corresponding Lagrangian is

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi, \tag{2.3}$$

where ψ is the Dirac spinor for the particle and $\bar{\psi}$ is the corresponding adjoint spinor, γ^μ are the Dirac gamma matrices and m is the mass of the particle. It is easy to see that 2.3 is invariant under $U(1)$ global phase transformations of the form

$$\psi(x) \rightarrow \psi'(x) = e^{i\alpha}\psi(x), \tag{2.4}$$

where α is a real constant. Applying local phase transformations of the form

$$\psi(x) \rightarrow \psi'(x) = e^{i\alpha(x)}\psi(x), \tag{2.5}$$

to Equation 2.3 yields the Lagrangian

$$\mathcal{L}'_{\text{Dirac}} = \mathcal{L}_{\text{Dirac}} - \bar{\psi}\gamma^\mu(\partial_\mu \alpha(x))\psi. \tag{2.6}$$

Thus, the Dirac Lagrangian is not invariant under local $U(1)$ phase transformations. This can be fixed by replacing the derivative ∂_μ with a *covariant derivative*

$$D_\mu \equiv \partial_\mu - ieA_\mu, \tag{2.7}$$

where e is the charge of the particle and A_μ is a vector field that transforms as

$$A_\mu \to A'_\mu = A_\mu + \frac{1}{e}\partial_\mu \alpha(x). \tag{2.8}$$

Thus the field A_μ couples to a Dirac particle of charge $-e$. This is the photon field of QED. The Dirac Lagrangian is now invariant under local phase transformations. However, kinetic energy and mass terms must be added if this is to represent a real field. Adding these to the Lagrangian gives

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m_A^2 A_\mu A^\mu, \tag{2.9}$$

where $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ is known as the *field strength tensor* and m_A is the mass of the photon. Under local phase transformations of the form given in Equation 2.5, the first term of Equation 2.9 has already been made invariant, the second is also invariant, however the third term is not. For this Lagrangian to be invariant under $U(1)$ local phase transformations, it is required that $m_A = 0$. As the photon is observed to be massless, this is not a problem. The QED Lagrangian is therefore

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + e\bar{\psi}\gamma^\mu \psi A_\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{2.10}$$

2.3 Quantum chromodynamics (QCD)

Quantum chromodynamics (QCD) describes the strong interaction between quarks and gluons and is based on the non-Abelian $SU(3)_C$ symmetry group. The description will follow the same as that used for QED, but using $SU(3)$ symmetries in the place of $U(1)$ due to the fact in QCD there are three colour charges as opposed to the single charge of QED. The free Lagrangian for quarks is

$$\mathcal{L}_{\text{free}} = \bar{q}_j(i\gamma^\mu \partial_\mu - m)q_j \text{ for } j = 1, 2, 3\ , \tag{2.11}$$

where each q_j is the spinor field for a colour j. In the case of $SU(3)$ symmetries, the Lagrangian is required to be invariant under transformations of the type

$$q(x) \rightarrow e^{i\alpha_k(x)T_k} q(x) \text{ for } k = 1, 2, ..., 8 \ , \tag{2.12}$$

where $\alpha_k(x)$ are the parameters of the group and T_k are a set of traceless, linearly independent, 3×3 matrices. As previously, a covariant derivative is introduced

$$D_\mu = \partial_\mu + igT_k G_\mu^k, \tag{2.13}$$

where g is the colour charge and G_μ^k is the gluon field for colour k. In order that local gauge invariance is satisfied, G_μ^k must transform as

$$G_\mu^k \rightarrow G_\mu'^k = G_\mu^k - \partial_\mu \alpha_k - g f_{ijk} \alpha_i G_\mu^j, \tag{2.14}$$

where f_{ijk} are the structure constants of the group defined by the commutator

$$[T_i, T_j] = i f_{ijk} T_k. \tag{2.15}$$

The field strength tensor for QCD takes the form

$$G_i^{\mu\nu} = \partial^\mu G_i^\nu - \partial^\nu G_i^\mu - g f_{ijk} G_j^\mu G_k^\nu. \tag{2.16}$$

The final term in Equation 2.16 is required by gauge invariance and accounts for gluon self interaction, i.e. requiring gauge invariance implies that gluons must self-interact, which is observed to be true in nature. Adding the kinetic energy term from the field strength tensor to Equation 2.11 gives the QCD Lagrangian

$$\mathcal{L}_{\text{QCD}} = \bar{q}(i\gamma^\mu \partial_\mu - m)q - g(\bar{q}\gamma^\mu T_i q)G_\mu^i - \frac{1}{4} G_{\mu\nu}^i G_i^{\mu\nu}. \tag{2.17}$$

Note that a term to account for the gluon mass can also be added as in Equation 2.9, but as previously, requiring gauge invariance implies that the gluon must be massless.

2.4 The weak interaction

The weak interaction is based on the non-Abelian $SU(2)_L$ symmetry group, with the L denoting that the corresponding vector field couples only to the left-handed chiral fermions and right-handed chiral antifermions. The left handed fermions are placed in

weak-isospin, $I_W = \frac{1}{2}$, doublets

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L, \begin{pmatrix} u \\ d' \end{pmatrix}_L, \begin{pmatrix} c \\ s' \end{pmatrix}_L, \begin{pmatrix} t \\ b' \end{pmatrix}_L, \tag{2.18}$$

where the quark states denoted q' refer to the weak eigenstates of the respective quarks. These eigenstates are mixtures of the mass eigenstates, with the mixing specified by the CKM matrix. The upper field of each doublet corresponds to the third component of weak-isospin, $I_W^3 = +1/2$ and the lower to $I_W^3 = -1/2$. The right-handed fermions are placed in weak-isospin singlets with $I_W = 0$

$$e_R^-, \mu_R^-, \tau_R^-, d_R, u_R, s_R, c_R, b_R, t_R. \tag{2.19}$$

The same procedure as in Section 2.3 can be followed using the $SU(2)$ symmetry producing 3 gauge fields, W_μ^1, W_μ^2 and W_μ^3. This can be used to obtain the locally gauge invariant Lagrangian for the weak interaction

$$\mathcal{L}_{\text{weak}} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - g_W(\bar{\psi} T_i \psi) W_\mu^i - \frac{1}{4} W_{\mu\nu}^i W_i^{\mu\nu}, \tag{2.20}$$

where $i = 1, 2, 3$, g_W is a constant, T_i are a set of 2×2, linearly independent, traceless matrices and

$$W_i^{\mu\nu} = \partial^\mu W_i^\nu - \partial^\nu W_i^\mu - g\varepsilon_{ijk} W_j^\mu W_k^\nu, \tag{2.21}$$

where ε_{ijk} is the Levi-Civita tensor. As before, adding a mass term to the Lagrangian would violate the local gauge symmetry. However, this poses a problem as the bosons corresponding to the weak interaction are known to be massive.

2.5 Spontaneous symmetry breaking and the Higgs mechanism

The problem of the masses of the weak bosons is solved by a process known as *spontaneous symmetry breaking*, the process by which massless particles can obtain mass. To demonstrate this using a simple example, consider a real scalar field, ϕ with

a Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - V(\phi), \tag{2.22}$$

$$V(\phi) = \frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4, \tag{2.23}$$

where μ and λ are the mass and self interaction terms respectively. In the nominal

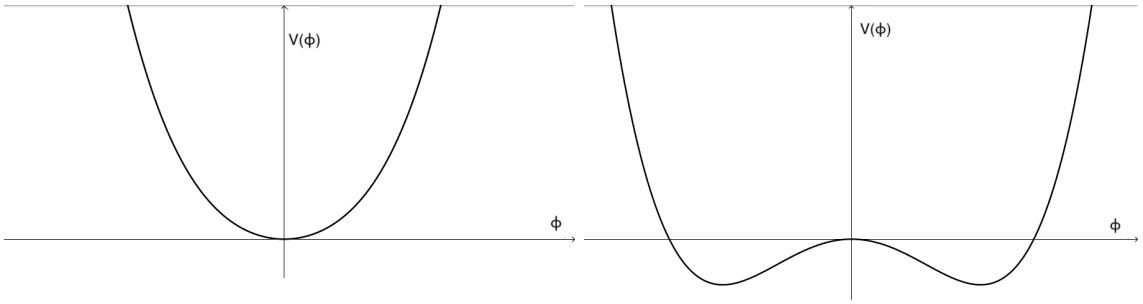


Fig. 2.1 The potential of Equation 2.23 for the cases $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right).

case where $\mu^2 > 0$, the potential is the shape given in Figure 2.1 (left), i.e. a scalar boson with mass μ. In the case where $\mu^2 < 0$, the minimum of the potential is no longer at $\phi = 0$ as shown in Figure 2.1 (right). There are now two possible vacuum states given by $\phi = \pm v$, spontaneously breaking the symmetry of the Lagrangian. The quantity v is known as the *vacuum expectation value*, and it is trivial to deduce that $v = \sqrt{-\mu^2/\lambda}$. The potential can now be expanded about one of the minima. Choosing $\phi = +v$, a scalar field η is introduced

$$\phi(x) = v + \eta(x). \tag{2.24}$$

The Lagrangian of Equation 2.22 therefore becomes

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \lambda v^2\eta^2 - \lambda v\eta^3 - \frac{1}{4}\lambda\eta^4 + \frac{1}{4}\lambda v^4. \tag{2.25}$$

The second term accounts for the mass of the boson associated with this field and the third and fourth terms, its self interaction. Thus a massive scalar field has been generated by the process of spontaneous symmetry breaking. To generalise to the case of a complex scalar field, now consider a complex scalar field of the form

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2), \tag{2.26}$$

which is described by the $U(1)$ globally gauge invariant Lagrangian

$$\mathcal{L} = (\partial_\mu \phi)^*(\partial^\mu \phi) - \mu^2 \phi^* \phi - \lambda(\phi^* \phi)^2. \tag{2.27}$$

The same procedure as above for the case $\mu^2 < 0$ can be performed here by expanding about the point $\phi_1 = v$, ϕ_2 by making the substitutions $\phi_1(x) = \eta(x) + v$ and $\phi_2(x) = \xi(x)$. The field can therefore be expressed as

$$\phi(x) = \frac{1}{\sqrt{2}}(v + \eta(x) + i\xi(x)). \tag{2.28}$$

Substituting this into the Lagrangian of Equation 2.27 yields

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) + \frac{1}{2}(\partial_\mu \xi)(\partial^\mu \xi) + \lambda \nu^2 \eta^2 - V_{\text{int}}(\eta, \xi), \tag{2.29}$$

where $V_{\text{int}}(\eta, \xi)$ contains all the interaction terms. This represents two scalar fields, one massive, η, with a mass of $v\sqrt{2\lambda}$ and one massless, ξ. Note that this Lagrangian is not invariant under a local phase transformation of the type given in Equation 2.5. As in the case of QED, a covariant derivative, D_μ, is introduced, defined as

$$D_\mu = \partial_\mu - ieA_\mu, \tag{2.30}$$

where the gauge field A_μ transforms as

$$A_\mu \to A'_\mu = A_\mu + \frac{1}{e}\partial_\mu \alpha. \tag{2.31}$$

Inserting these definitions into Equation 2.29 and adding the kinetic energy term for the field A_μ yields

$$\mathcal{L} = \underbrace{\frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \lambda v^2 \eta^2}_{\text{massive } \eta \text{ boson}} + \underbrace{\frac{1}{2}(\partial_\mu \xi)(\partial^\mu \xi)}_{\text{massless } \xi \text{ boson}} - \underbrace{\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}e^2 v^2 A_\mu A^\mu}_{\text{massive gauge boson}} - V_{\text{int}}(\eta, \xi) - evA_\mu \partial^\mu \xi. \tag{2.32}$$

Thus, from adding a scalar field with a spontaneously broken symmetry, the boson corresponding to the gauge field A_μ can have a mass without breaking local gauge invariance. In doing so, a massive η boson and a massless ξ boson, known as a *Goldstone boson* have been created. The final term in Equation 2.32, describing the interaction between the gauge field and the Goldstone boson is unphysical. This term can be

removed by making the gauge transformation

$$A_\mu \to A'_\mu = A_\mu + \frac{1}{ev}\partial_\mu \xi. \tag{2.33}$$

Applying this change to Equation 2.32 and expanding the interaction terms results in the Lagrangian

$$\begin{aligned}\mathcal{L} = &\underbrace{\frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \lambda v^2 \eta^2}_{\text{massive } \eta \text{ boson}} - \underbrace{\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}e^2 v^2 A_\mu A^\mu}_{\text{massive gauge boson}} \\ &\underbrace{- e^2 v A_\mu A^\mu \eta + \frac{1}{2}e^2 A_\mu A^\mu \eta^2}_{A\text{-}\eta \text{ interactions}} \underbrace{- \lambda v \eta^3 - \frac{1}{4}\lambda \eta^4}_{\eta \text{ self interaction}}.\end{aligned} \tag{2.34}$$

Note that this result is the equivalent of taking Equation 2.28 to be $\phi = \frac{1}{\sqrt{2}}(v + \eta(x))$. To summarise, by introducing a complex scalar field with a spontaneously broken symmetry that couples to a locally gauge invariant field, A_μ, the mass for the gauge boson is generated along with a massive boson for the scalar field, η. These bosons have masses of ev and $\sqrt{2\lambda}v$ respectively. This is known as the *Higgs mechanism*, with the massive boson, η, being analogous to the Higgs boson.

2.6 Electroweak symmetry breaking in the Standard Model

To generate mass terms for the bosons of the weak interaction, the symmetry $U(1)_Y$ must first be introduced. This is the $U(1)_Q$ symmetry of QED, replacing charge Q with the quantity Y known as the *weak hypercharge*. This is defined as

$$Y = 2(Q - I^3_W). \tag{2.35}$$

This symmetry has a corresponding field B_μ which behaves in an analogous way to the A_μ field of QED. The Standard Model Higgs boson is generated by the breaking of the $U(1)_Y \times SU(2)_L$ symmetry, known as the *electroweak* symmetry, in the manner described in the previous section. The entire procedure for the $U(1)_Y \times SU(2)_L$ case will not be shown in the same detail given above, but a summary is as follows. A

doublet of complex scalar fields is introduced

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{2.36}$$

the Lagrangian for this field is given by

$$\mathcal{L} = (\partial^\mu \Phi)^\dagger (\partial_\mu \Phi) - V(\Phi), \tag{2.37}$$

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2. \tag{2.38}$$

Φ is expanded about the vacuum expectation value, v as

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \tag{2.39}$$

where $H(x)$ is a real scalar field, namely the Standard Model Higgs field. The covariant derivative for the $U(1)_Y \times SU(2)_L$ symmetry is

$$D_\mu = \partial_\mu + i g_W T_k W_\mu^k + i g' \frac{Y}{2} B_\mu. \tag{2.40}$$

Expanding the kinetic term of Equation 2.37 yields

$$|D_\mu \Phi|^2 = \frac{1}{2}(\partial_\mu H)^2 + \frac{1}{8} g_W^2 (v+H)^2 \left|W_\mu^1 + i W_\mu^2\right|^2 + \frac{1}{8}(v+H)^2 \left|g_W W_\mu^3 - g' B_\mu\right|^2. \tag{2.41}$$

From this, the mass terms are mixtures of W_μ^i and B_μ and thus the W_μ^i and B_μ do not correspond to the physical fields. These are obtained through the linear combinations

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp i W_\mu^2), \tag{2.42}$$

$$Z_\mu = \frac{W_\mu^3 g_W - g' B_\mu}{\sqrt{g_W^2 + g'^2}} \tag{2.43}$$

and

$$A_\mu = \frac{W_\mu^3 g_W + g' B_\mu}{\sqrt{g_W^2 + g'^2}}, \tag{2.44}$$

where $W_\mu^\pm$ and Z_μ are the fields giving rise to the massive W and Z bosons respectively and A_μ is the massless photon field of QED.

From expanding the terms of Equation 2.41, the masses of the bosons can be identified as

$$\begin{aligned} m_{W^\pm} &= \frac{1}{2} g_W v, \\ m_Z &= \frac{1}{2} v \sqrt{g_W^2 + g'^2}, \\ m_A &= 0, \\ m_{\mathrm{H}} &= v\sqrt{2\lambda}. \end{aligned} \tag{2.45}$$

The Higgs mechanism applied to the $U(1)_Y \times SU(2)_L$ symmetry successfully generates masses for the W and Z bosons as well as leaving the photon massless. It also predicts the existence of a massive scalar boson, known as the Higgs boson, the existence of which was verified by the ATLAS and CMS collaborations in 2012 [4, 5].

As well as generating the masses of the gauge bosons, the Higgs mechanism is also able to generate the masses of the fermions via the addition of *Yukawa terms* to the relevant Lagrangians. As with the cases of the gauge bosons, adding a term

$$-m\bar{\psi}\psi = -m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R), \tag{2.46}$$

to a Lagrangian violates the local gauge invariance. To account for this, a gauge invariant term is introduced, which is for the case of the $SU(2)$ e-ν_e doublet

$$\mathcal{L}_e = -G_e \left[(\bar{\nu}_e \ \bar{e})_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \bar{e}_R (\phi^{+*} \ \phi^{0*}) \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \right], \tag{2.47}$$

where G_e is a constant known as a *Yukawa coupling constant*. Using the choice of Higgs potential given in Equation 2.39, Equation 2.47 becomes

$$\mathcal{L}_e = -\frac{G_e}{\sqrt{2}} v(\bar{e}_L e_R + \bar{e}_R e_L) - \frac{G_e}{\sqrt{2}} (\bar{e}_L e_R + \bar{e}_R e_L) H, \tag{2.48}$$

from which the electron mass can be identified as

$$m_e = \frac{G_e v}{\sqrt{2}}. \tag{2.49}$$

Thus the Higgs mechanism can also generate masses for the fermions. It is also noteworthy from the second term of Equation 2.48, that the strength of the coupling between the Higgs field and the electron field is in direct proportion to the mass of the electron. This same procedure can be repeated to also generate masses for the quarks.

It should also be noted in Equation 2.45 that m_{H}, being the only mass which depends on λ, is a free parameter in the Standard Model. However, whilst a measurement of m_{H} is not a direct test of SM predictions, the production cross-sections of the Higgs boson, as well as its branching ratios to various decay modes depend on m_{H}. So by measuring mass of the Higgs boson, as well as its production cross sections and branching ratios, a test of the SM can be performed.

2.7 Production of Higgs bosons in hadronic collisions

In a hadronic collider, such as the LHC, the production of Higgs bosons must be initiated by interactions between quarks or gluons. As mentioned previously, the Higgs boson couples more strongly to more massive particles, so dominant production modes will involve heavier particles. The four most dominant production modes for Standard Model Higgs bosons are as follows

Gluon-gluon fusion (ggF) Whilst the gluon, being massless, does not directly interact with the Higgs boson, two gluons can interact via a virtual top quark loop, with the massive top quarks sharing a vertex with a Higgs boson, as shown in Figure 2.2a. This is the dominant production mode. Higgs boson production via this process is an order of magnitude more common than the next most dominant production mode.

Vector boson fusion (VBF) This is the next most dominant production mode, occurring when two W or Z bosons are radiated by quarks in a collision and then fuse to create a Higgs, shown in Figure 2.2b.

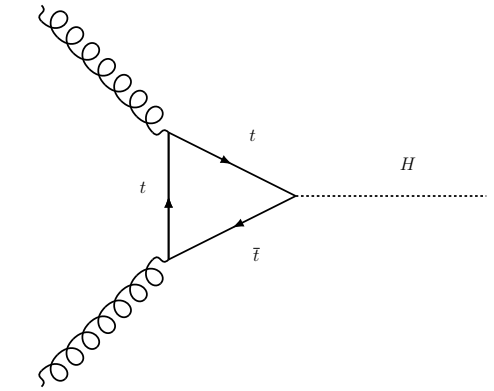
Vector boson associated production (VH) A pair of quarks can fuse to create a W or Z boson, which can then radiate a Higgs boson, Figure 2.2c.

$t\bar{t}H$ **and** $b\bar{b}H$ **production** Similarly to the VH case, top and bottom quark pairs can also radiate a Higgs boson, see Figure 2.2e

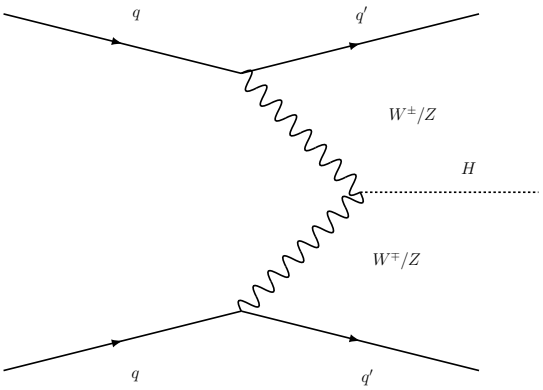
tH **production** This is the rarest production mode considered, shown in Figure 2.2d. This occurs when a single top quark radiates a Higgs boson. It is an extremely rare process, occurring 100 times less often than ggF.

The cross sections for each production mode in pp collisions for a centre of mass energy of $\sqrt{s} = 13$ TeV are shown in Figure 2.3. These are calculated using parton distribution functions and perturbative calculations on the strong and electroweak

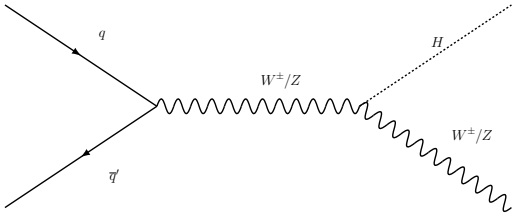
couplings. NLO refers to these calculations being *next to leading order*, NNLO *next to next to leading order* and N3LO being one order higher NNLO. As shown, the ggF process (denoted $pp \to H$ here) is far more common than the next most common mode, VBF (denoted $pp \to qqH$).



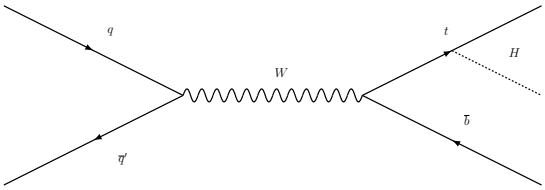
(a) Gluon-gluon fusion (ggF)



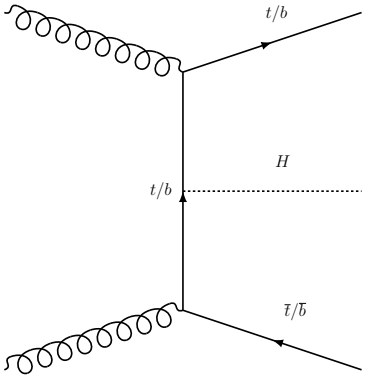
(b) Vector boson fusion (VBF)



(c) Vector boson associated production (VH)



(d) tH production



(e) $t\bar{t}H/b\bar{b}H$ production

Fig. 2.2 The dominant processes for Higgs boson production in proton-proton collisions.

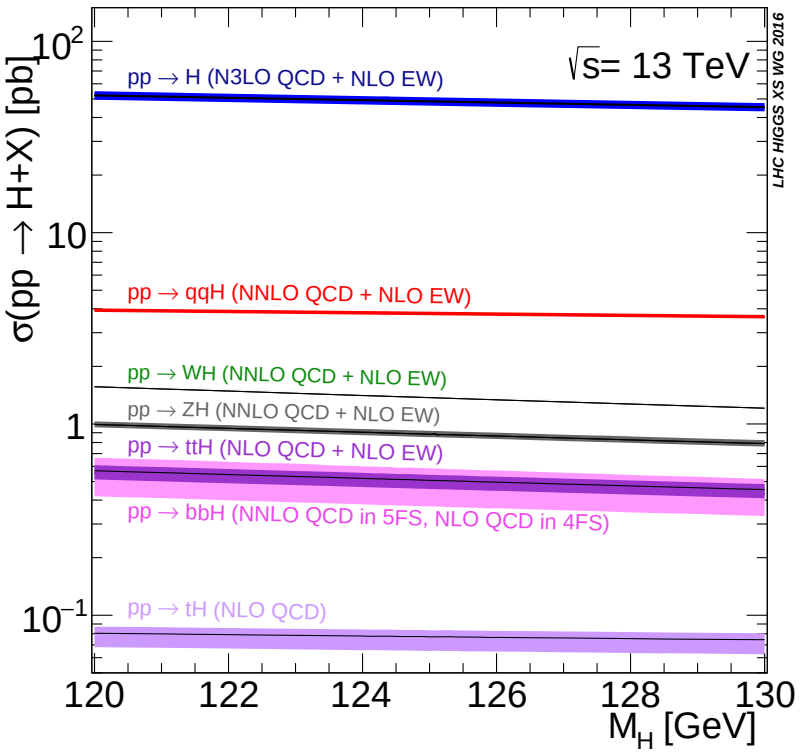


Fig. 2.3 Higgs boson production cross section vs m_{H} for pp collisions with a centre of mass energy of 13 TeV. [15]

2.8 Four-lepton decays of Higgs bosons

Due to the Higgs coupling to mass, it will favour decays via heavier particles, provided the combined mass of the decay products is $< m_{\mathrm{H}}$. Figure 2.4 shows the branching ratio for the leading decay modes of the Higgs vs its mass. Note that the decays to the massless photons and gluons are allowed as these can happen via top quark loops, i.e. the reverse of Figure 2.2a. Also note that decays to heavy gauge bosons are allowed despite the fact that $m_{\mathrm{H}} < 2m_W, 2m_Z$ as one of the W's or Z's is produced off-shell. Clearly, the most probable decay mode is $H \to b\bar{b}$. However, this is not the easiest channel to use for performing a mass measurement for two reasons. Firstly b quarks produced hadronise, radiating gluons, which themselves generate more hadrons creating hadronic "jets" which are difficult to resolve accurately in a detector. Secondly, this process has a large background from $W+$ jets and top quark processes. The next most dominant decay mode is $H \to W^+W^-$. Whilst the W's can decay without producing quarks, leptonic decays of W's produce neutrinos, which escape the detector without detection. The decay $H \to \tau^+\tau^-$ is also unsuitable, as τ's have a lifetime of $\sim 10^{-13}$

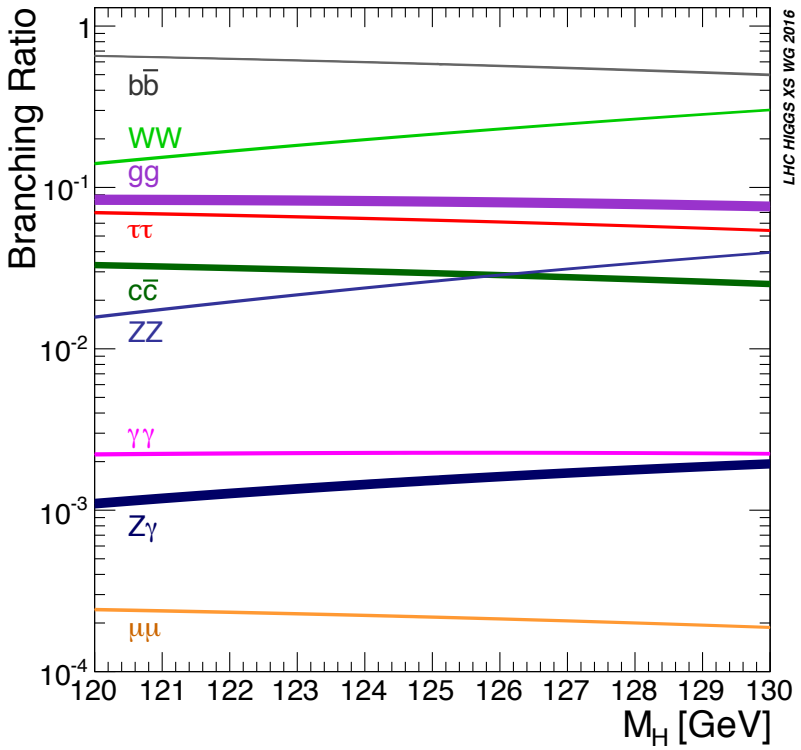


Fig. 2.4 SM Higgs boson branching ratio of various decay modes vs m_{H}. [15]

they decay before interacting with the detector, producing neutrinos as well as jets in most instances. The only remaining decay modes in order of branching ratio at $m_{\mathrm{H}} = 125$ GeV are $H \to ZZ$, $H \to \gamma\gamma$, $H \to Z\gamma$ and $H \to \mu\mu$. As Z bosons can decay into electron/muon pairs, these can all produce final states containing only photons, muons and electrons which are well reconstructed by ATLAS. The latter two however are unsuitable as these decays are very rare, have a large background and to date have not been observed. Although the branching ratio of the Higgs decay to two Z bosons is much larger than that of the decay to two photons, the branching ratio for the decay to two Z's where these then decay to either electrons or muons is around 0.1 that of $H \to \gamma\gamma$. Although, this decay is relatively rare, it has a small background associated with it, which makes it a suitable channel for which to perform a measurement of m_{H}. The Feynman diagram for this decay is shown in Figure 2.5. $H \to \gamma\gamma$ can also be used to perform a mass measurement. Although the number of events expected is larger than that of $H \to ZZ^* \to 4\ell$, the background contribution is larger and photons are less precisely resolved in ATLAS than muons. For this reason, the $H \to ZZ^* \to 4\ell$ channel provides the most precise measurement of m_{H}.

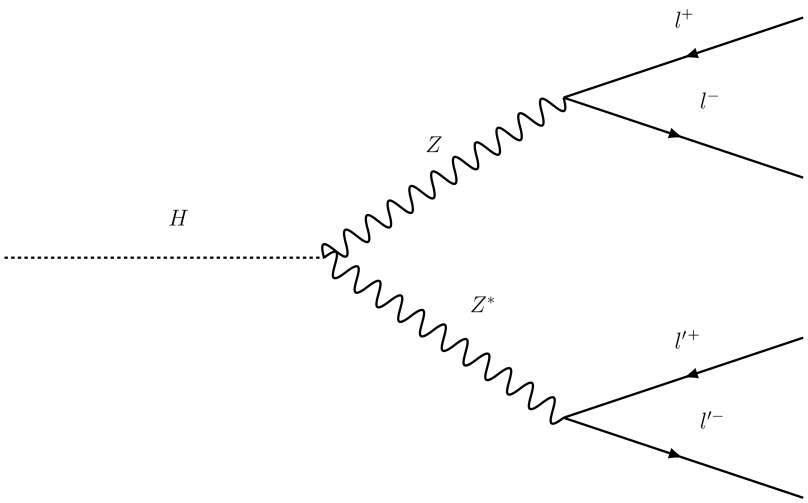


Fig. 2.5 The Feynman diagram for the Higgs decay to four leptons ($H \to ZZ^* \to 4\ell$) (for $l = e, \mu$). The "*" indicates the off-shell Z.

Chapter 3

The Large Hadron Collider and the ATLAS experiment

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) at the European Organisation for Nuclear Research (CERN) is currently the largest particle accelerator in operation. The primary purpose for its construction was the discovery of the Higgs boson, which was achieved in 2012. It is currently being used to study the Higgs boson as well as conduct precision tests on the SM and search for new physics. It is a circular collider, designed to collide bunches of 10^{11} protons 40 million times per second at centre of mass energies up to $\sqrt{s} = 14$ TeV and an instantaneous luminosity of 10^{34} cm^{-2} s^{-1} [16]. At four points around the circumference of the LHC, the protons are collided. The ATLAS experiment is located at one of these points and is the experiment used to collect the data presented here.

3.2 The ATLAS detector

The ATLAS experiment is a general purpose detector optimised to detect previously unseen particles directly and perform precision tests on the Standard Model [17]. ATLAS uses a cylindrical coordinate system, with the nominal interaction point as the origin with the z direction being that of the beam-pipe. The $x - y$ is transverse to the z axis, the x axis points from the interaction point to the centre of the LHC ring and y axis points upwards. The azimuthal angle, ϕ, is measured around the beam axis and

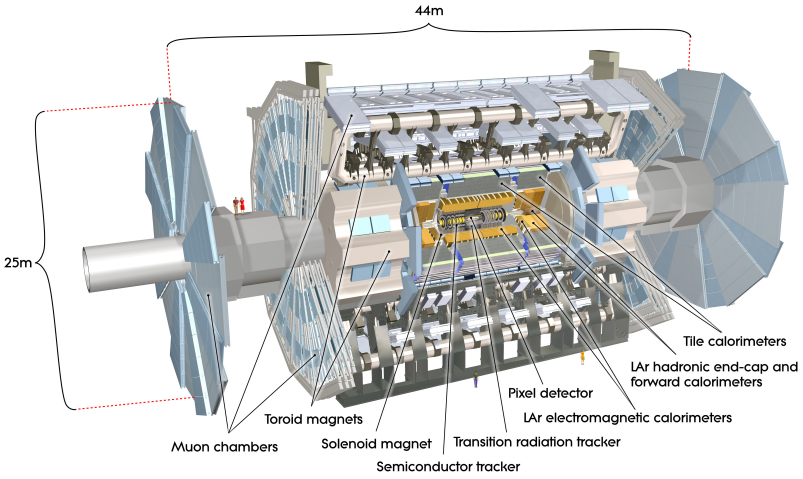


Fig. 3.1 Cutaway diagram of the ATLAS detector. [17]

the polar angle, θ, is measured from the beam axis. Pseudorapidity, η, is defined as

$$\eta = -\ln \tan \frac{\theta}{2}. \tag{3.1}$$

The angular separation between two objects is given by

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}, \tag{3.2}$$

where $\Delta\eta$ and $\Delta\phi$ are the differences in η and ϕ between the two objects. The detector is constructed in a concentric fashion with the LHC beam-pipe in the centre, providing full coverage in the azimuthal ϕ direction and in the pseudorapidity range $|\eta| < 2.5$. A diagram of the detector is shown in Figure 3.1. The innermost component, surrounding the beam-pipe, is the tracker system, used to measure the transverse momenta (p_{T}) of charged particles by tracking their path in a 2 T solenoidal magnetic field. Surrounding this are the electromagnetic (EM) and hadronic calorimeters. The former is used to measure the energies of electrons and photons and the latter, the energy of hadrons. The outermost component is the muon system, used to determine the transverse momenta of muons by tracking their paths in a toroidal magnetic field. ATLAS also relies on a trigger system to select, in real time, events of interest.

3.2.1 Inner detector

As shown in Figure 3.2, the ATLAS tracking system is composed of three components. As with the other parts of the detector, these are arranged in a concentric fashion.

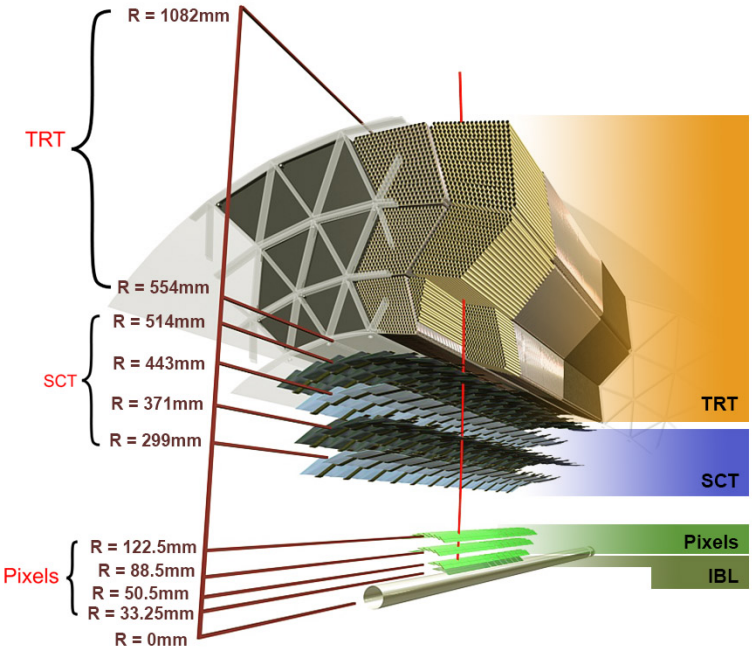


Fig. 3.2 Diagram showing the elements of ATLAS tracking system. [18]

From inside to outside, these are the silicon pixel trackers (pixel), silicon microstrip trackers (SCT) and transition radiation trackers (TRT). The pixel consists of sensors of size $50{\times}400$ μm^2 and the SCT of stereo pairs of silicon microstrips, with one microstrip in each pair parallel to the beam-pipe and the other at an angle of 40 mrad. These two components provide the highest resolution tracking information in the range $|\eta| < 2.5$. The TRT consists of straw tubes filled with a gaseous $Xe/CO_2/O_2$ mixture and provide tracking information in $r - \phi$ space within a pseudorapidity coverage of $|\eta| < 2.0$. A track in this range with $p_T > 0.5$ GeV will cross a minimum of 36 straws, the exception being in the region $0.8 < |\eta| < 1.0$ where the minimum number of straws crossed can be as low as 22. These components are ordered by the granularity they provide with the pixel detector providing the highest. This ordering reflects the fact that the density of particle tracks is higher nearer the interaction point, thus a higher granularity is required. The inner dector is immersed in a 2 T solenoidal magnetic field, deflecting the charged particles, allowing their transverse momentum to be measured from the curvature of their tracks.

3.2.2 Calorimetry

Electromagnetic showers and calorimetry

To measure the energies of electrons and photons at GeV scales, their behaviour must be understood at the sub-MeV scale. The measurement of electron/photon energies is performed by measuring the products of *electromagnetic showers* (EM showers). In the case of electrons, the process of showering starts with the electron interacting with the nuclei of the detector material and radiating a photon in a process known as *Bremsstrahlung* [19]. For photons, an interaction with a nucleus can cause the photon to produce an electron-positron pair in a process known as *pair production*. Subsequent electrons (photons) produced then proceed to radiate photons (pair produce) themselves causing a "cascade" of electrons and photons. A diagram of the beginning

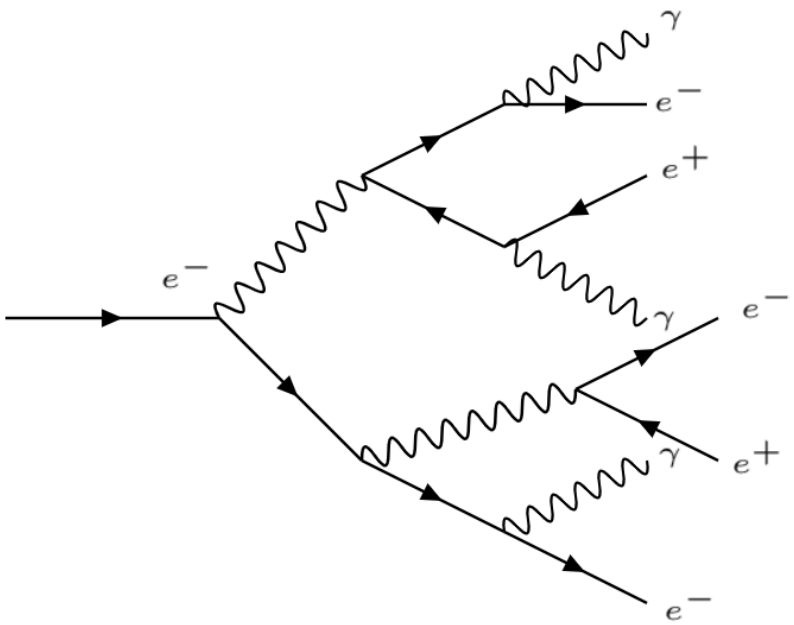


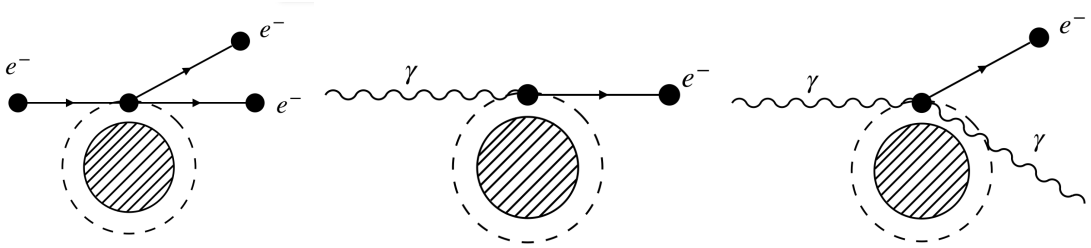
Fig. 3.3 Schematic diagram of an electromagnetic shower initiated by an electron. The incident electron radiates a Bremsstrahlung photon, which itself creates an electron positron pair through the process of pair production. Note that each of these vertices is initiated by the exchange of a photon with an atom of the detector material, which is not shown.

of an electromagnetic shower from an electron is shown in Figure 3.3. The processes of Bremsstrahlung and pair production are the dominant ways in which electrons and photons respectively interact with the detector material at high energies. At lower energies, electrons may also interact with the detector via ionisation (Figure 3.4a).

The critical energy at which Bremsstrahlung and ionisation occur at an equal rate is given by

$$\epsilon_c = \frac{610 \text{ MeV}}{Z + 1.24}, \tag{3.3}$$

where Z is the atomic number of the material [19]. The dominant process by which photons interact at low energies is the photoelectric effect (Figure 3.4b) and by Compton scattering at intermediate energies (Figure 3.4c). The cross sections in lead for each of



(a) Ionisation (b) The photoelectric effect (c) Compton scattering

Fig. 3.4 The three dominant processes by which electrons and photons lose energy at medium to low energies ($\lessapprox$ 10 MeV).

these processes vs photon energy are shown in Figure 3.5. Here, the photoelectric effect is the dominant process up to approximately 1 MeV, where Compton scattering becomes dominant. Above these energies, pair production dominates. It is also important to note that the angular distribution for the former two processes is approximately isotropic, whereas that of pair production is highly directional. This has consequences in the design of calorimeters as these low energy particles are the ones which cause the signal in the detector. As these happen isotropically, structures other than the traditional "sandwich" can be used. As a shower progresses through a material it will tend to spread laterally. At the start of an electron shower, this will be due to multiple scattering of electrons, as these will move away from the shower axis. Near the end of the shower in the lower energy regime, this will tend to be due to the isotropic nature of the processes producing lower energy electrons and photons. The longitudinal depth a shower penetrates and its lateral spread depend on the material in which the showering occurs. For convenience, these are parameterised in terms of two approximately material-independent quantities. These are the *radiation length* (X_0), which describes the longitudinal depth of the shower, and the *Molière radius* (ρ_M), which describes the lateral spread of the shower. X_0 is the distance over which an electron will on average lose $1 - e^{-1}$ of its original energy. ρ_M is proportional to the ratio of X_0 with the critical energy. ρ_M scales with A/Z, therefore its value does not change significantly between different materials. However, X_0 scales with A/Z^2, so

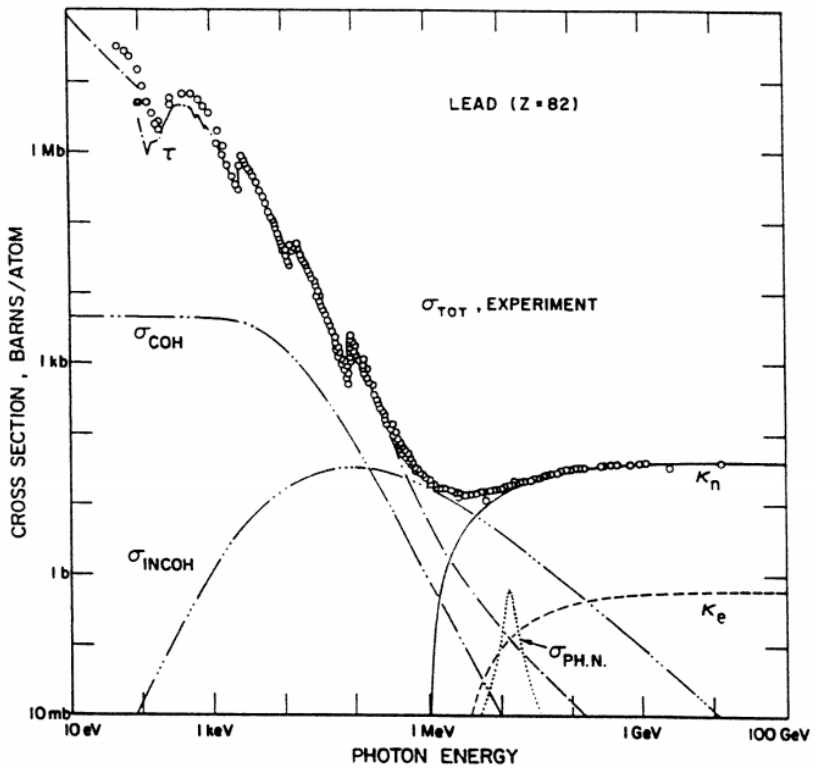
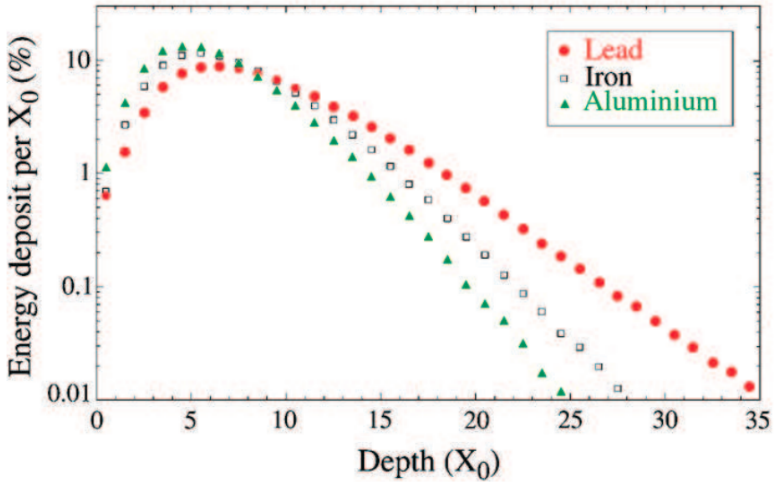
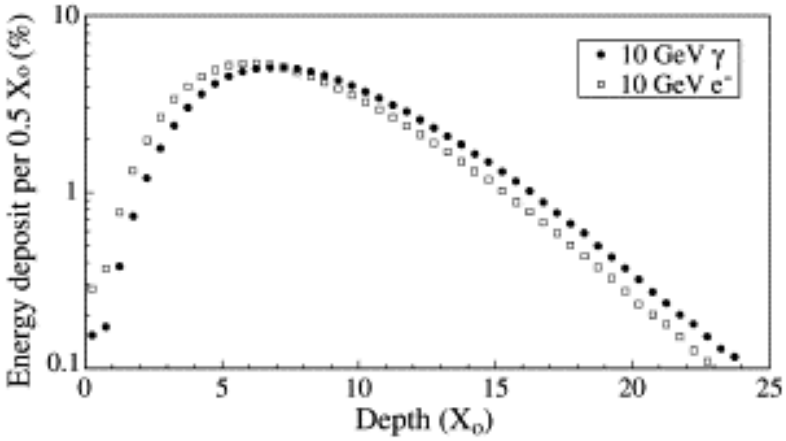


Fig. 3.5 The cross sections for each of the processes that contribute to photon scattering vs incident photon energy. The dominant process at low energies is the photoelectric effect, the cross section for which is denoted τ here. The Compton effect dominates at energies of ~ 1 MeV, the cross section for which is denoted σ_{INCOH} . Above these energies, pair production from interactions with the atomic nuclei dominates, κ_{n}. The cross sections for sub-leading processes of Rayleigh scattering (σ_{COH}), pair production from atomic electron interaction (κ_{e}) and nuclear photoabsorption ($\sigma_{\mathrm{PH.N}}$) are also shown along with the total cross section from experiment (σ_{TOT}). [20]

has some material dependence. The parameterisation of showers in terms of X_0 also has some dependence on the energy of the initial photon/electron. Figure 3.6a shows the differences between the longitudinal shower profiles in lead, iron and aluminium. It takes more radiation lengths to contain an electron shower in a lead absorber than an iron or aluminium one. Whilst it takes more radiation lengths of lead to contain an electron shower than that of iron, it should be noted that this corresponds to a smaller volume of lead, as the radiation length of lead is less than half that of iron [13].



(a)



(b)

Fig. 3.6 The longitudinal difference between electron initiated showers in different materials (a) and photon and electron initiated showers in tin (b). The fraction of initial energy deposited per X_0 is shown for a 10 GeV electron in lead, iron and aluminium absorbers (a) and per 0.5 X_0 for 10 GeV photons and electrons in a tin absorber (b). These were both calculated from Monte Carlo simulations. [19, 21]

Showers initiated by photons and electrons develop differently from one another due to the fact that electrons tend to radiate photons sooner than a photon tends to pair produce. As Figure 3.6b shows, showers initiated by photons tend to deposit more of their energy further into the absorber than those initiated by electrons.

ATLAS electromagnetic calorimeter

The ATLAS electromagnetic calorimeter contains 22 X_0 of lead absorber [22], which from Figure 3.6a will absorb approximately 99% of the energy deposited by a 10 GeV electron. As discussed previously, the amount contained will be lower for photons as well as for higher energy electrons. However, as Figure 3.6a shows, extending the calorimeter further outward results in diminishing returns for each additional radiation length of lead, so these losses are accounted for in the calibration (discussed in Section 4.2).

The ATLAS electromagnetic calorimeter system is a *sampling calorimeter*, meaning the absorber and active layers are distinct. A lead absorber is used along with an active layer of liquid argon (LAr). To detect a signal, copper electrodes are located in the middle of the LAr active layers with a high voltage system generating an electric field of around 1 kV/mm between the absorbers and electrodes [23]. The low energy electrons from the processes described above drift the gap and generate a signal. A diagram of the layout of the ATLAS calorimeters is shown in Figure 3.7. An "accordion" geometry is used for the calorimeter, as shown in Figure 3.8, as this provides full ϕ coverage with no gaps [24]. The LAr barrel calorimeter covers the range $|\eta| < 1.475$ and the two end-cap LAr calorimeters cover $1.375 < |\eta| < 3.2$. The barrel and endcap components are housed in separate cryostats. Additionally, in the region $|\eta| < 1.8$, a pre-sampler is used to correct for losses upstream. This correction is performed by considering the amount of energy deposited in the pre-sampler relative to the rest of the calorimeter. A larger deposit in the pre-sampler indicates that a larger amount of energy has been deposited upstream, so a larger correction must be applied to the calibrated energy. In the region $1.2 < |\eta| < 1.6$ there is a cryostat scintillator to correct for energy losses from the outside of the barrel cryostat. As im the case of the pre-sampler, a larger deposit implies a larger correction to the calibrated energy. A diagram of a segment of the LAr barrel calorimeter is shown in Figure 3.8. It is divided into three layers, the innermost having the finest granularity in η and the outermost having the lowest. As the second layer contains ~ 2.5 times more radiation lengths of lead than the other two layers combined, most of the energy deposition by electrons and photons happens here.

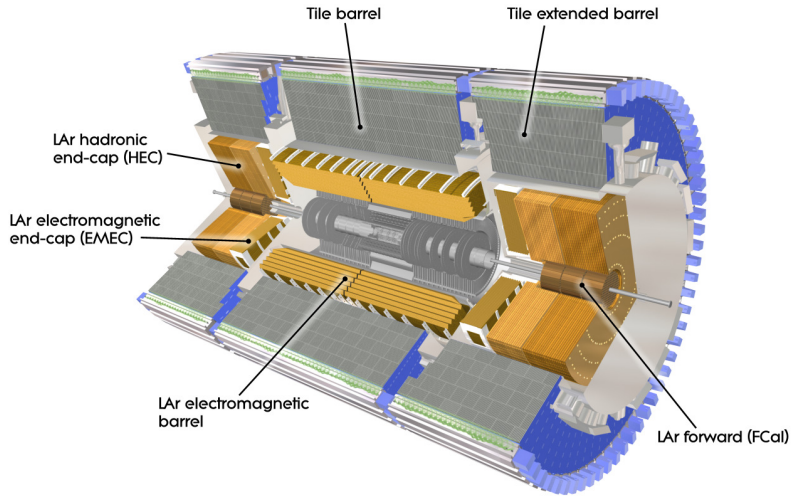


Fig. 3.7 A diagram of the layout of the ATLAS calorimeters. [17]

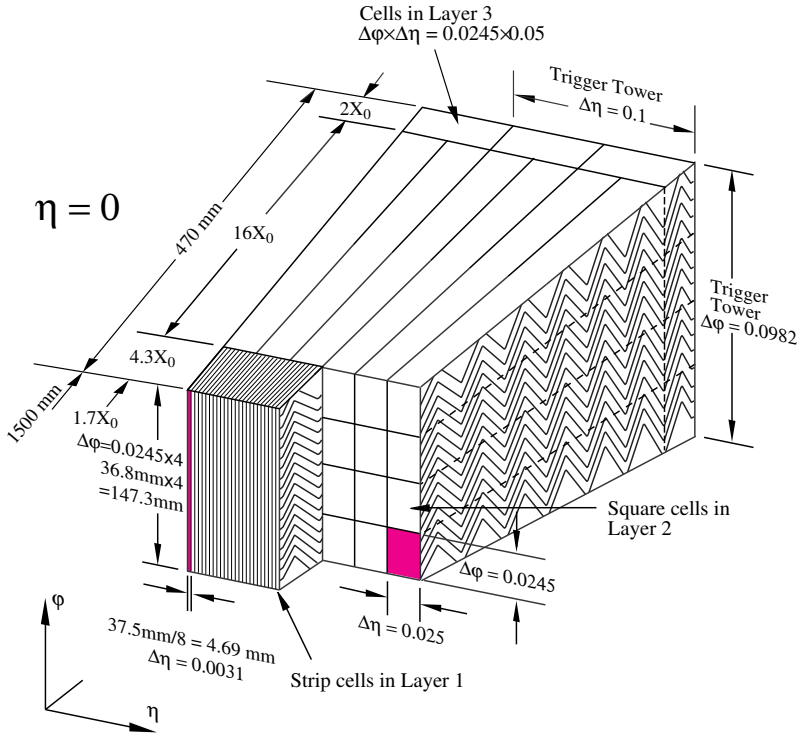


Fig. 3.8 A diagram of one of the barrel modules of the ATLAS electromagnetic calorimeter. [17]

Hadronic calorimeters

The hadronic calorimeters consist of a tile barrel calorimeter which covers the region $|\eta| < 1.0$, a tile extended barrel calorimeter covering $0.8 < |\eta| < 1.7$, a LAr hadronic end-cap calorimeter (HEC) covering $1.5 < |\eta| < 3.2$ and a forward LAr end-cap (FCAL)

which covers $3.1 < |\eta| < 4.9$. Their layout is shown in Figure 3.7. The tile calorimeters are both sampling calorimeters using a steel absorber and scintillating tiles as the active component. The HEC uses a copper absorber and a LAr active medium. The FCAL consists of three absorption layers, the innermost of which uses a copper absorber, optimised for EM interactions, the other two using a tungsten absorber, optimised for hadronic interactions with all three layers using LAr as the active component.

3.2.3 Muon system

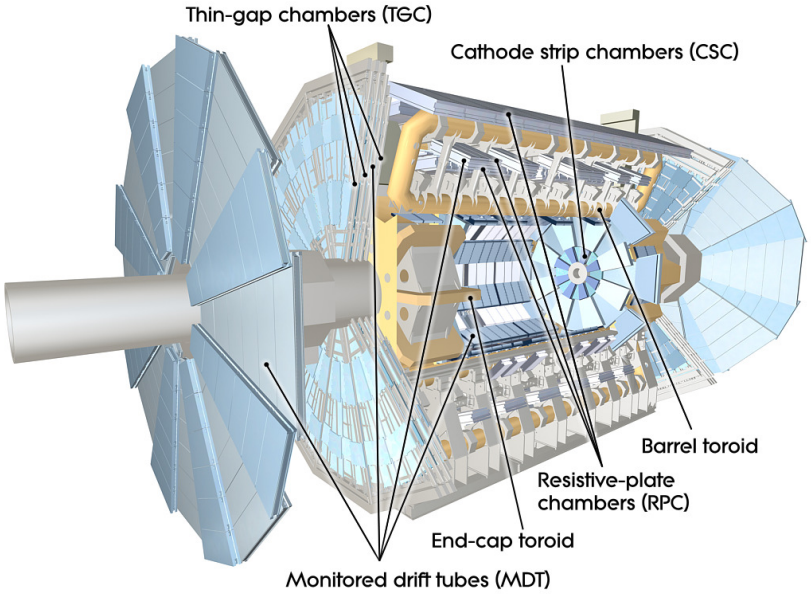


Fig. 3.9 A cutaway diagram of the ATLAS muon system. [17]

A diagram of the ATLAS muon system is shown in Figure 3.9. The magnetic field for the system is provided by a set of superconducting toroidal magnets in the region $|\eta| < 1.4$, by a set of end-cap toroid magnets for $1.6 < |\eta| < 2.7$ and by a combination of the two for $1.4 < |\eta| < 1.6$. Muon detection is provided by four separate systems. These are: monitored drift tubes (MDT) which cover $|\eta| < 2.7$, cathode strip chambers (CSC) for $2.0 < |\eta| < 2.7$, resistive plate chambers (RPC) for $|\eta| < 1.05$ and thin gap chambers (TGC) for $1.05 < |\eta| < 2.7$.

3.2.4 Trigger system

The event rate in ATLAS is ∼40 MHz, but it is not possible to record all events, so this must be reduced to a manageable rate. ATLAS uses a three-level trigger system

to select events of interest [25]. These are the Level-1 (L1), Level-2 (L2) and event filter. The first is a hardware-based triggering system and latter two are collectively referred to as the High-Level Trigger (HLT).

The majority of the reduction is performed by the L1, which reduces the event rate down to 100 kHz [26]. It does this by searching for high-p_{T} objects from collisions, i.e. photons, electrons, high-p_{T} jets, high-p_{T} muons, hadronically-decaying τ's and events with a large missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$). At this stage, the full detector granularity is not used. The location of each of the identified objects in $\eta - \phi$ space is used to determine regions of interest (RoIs), which are used at the next stage.

The L2 uses the full detector granularity, along with the RoI information, to select the events of potential interest. Events which pass the L2 are processed by the event filter, which performs its selections offline. At this stage, the entire event is reconstructed in order to make selections. The L2 and event filter collectively reduce the event rate down to 1 kHz. All of the events passing this stage are permanently stored at the CERN computing facilities.

Chapter 4

Electron and photon reconstruction and calibration

This chapter describes the ATLAS electron and photon reconstruction and calibration procedures as well as work by the author on a cross check to the photon calibration procedure and a study to investigate the effect of misidentifying photon conversions.

The ATLAS electron identification procedure is also described and work by the author in the correction of electron shower shapes is presented. An additional study by the author to correction the shower shapes using a method of Gaussian convolution is also described.

4.1 Electron and photon reconstruction

In ATLAS, a reconstructed electron is defined as an EM calorimeter cluster matched to an inner detector track. In the reconstruction of photons, it is necessary to define two types of object, depending on whether the photon converts to an electron-positron pair before reaching the calorimeter. These two different cases are known as *converted* and *unconverted photons*. The definition of a converted photon is an EM calorimeter cluster matched to a conversion vertex in the tracker and an unconverted photon as an EM calorimeter cluster not matched to a track or conversion vertex.

The reconstruction algorithm for electrons and photons proceeds in a three step process, namely:

1. Identify calorimeter clusters.

2. Reconstruct inner detector tracks.

3. Match tracks to clusters.

4.1.1 EM cluster reconstruction

Clusters are reconstructed in both the EM and hadronic calorimeters using a system of dynamically-sized topological clusters, abbreviated *topoclusters* [27]. These are formed using a 4-2-0 algorithm. That is, to begin forming a cluster, a seed is needed. This is a cell of the calorimeter for which $\left|\zeta_{\text{cell}}^{\text{EM}}\right| \geq 4$, where $\zeta_{\text{cell}}^{\text{EM}}$ is the cell significance. This is defined as

$$\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise, cell}}^{\text{EM}}}, \tag{4.1}$$

where $E_{\text{cell}}^{\text{EM}}$ is the energy of a given cell at the EM scale and $\sigma_{\text{noise, cell}}^{\text{EM}}$ is the expected cell noise. This expected noise accounts for electronic noise and an effect referred to as *pile-up* [28]. The latter refers to two effects. The first is *in-time* pile-up which is caused by multiple proton interactions in the same bunch crossing. The second is *out-of-time* pile-up which is caused by signals from other bunch crossings occurring shortly before or after the collision of interest, affecting the signal during the time for which the components of the detector are sensitive. This seed cell must not be in either the first layer of the EM calorimeter or the pre-sampler. All the cells neighbouring the seed cell are then added to this protocluster if they have $\left|\zeta_{\text{cell}}^{\text{EM}}\right| \geq 2$. This process then proceeds in an iterative fashion, with each cell added to the protocluster becoming a seed for the next iteration. If two protoclusters share a cell, they are merged. When all the cells satisfying this requirement have been added, all cells neighbouring the protocluster are added, without a significance requirement. Clusters containing at least two local maxima are then split into separate clusters. The definition of a local maximum is a cell for which $E_{\text{cell}}^{\text{EM}} > 500$ MeV and has at least four neighbours, none of which has a larger $E_{\text{cell}}^{\text{EM}}$ [29].

The EM cluster reconstruction proceeds by first applying a pre-selection to select the relevant topoclusters. This is done by first defining the EM energy of a cluster to be the energy deposited in the cells of the EM calorimeter. For clusters in the region $1.37 < |\eta| < 1.63$, the energy deposited in the pre-sampler and the cryostat scintillator is also added. From this, the electromagnetic fraction f_{EM} is defined as the ratio of the EM energy of a cluster to its total energy. Clusters are selected if they have $f_{\text{EM}} > 0.5$. This removes approximately 60% of clusters caused by pile-up without removing any signal clusters. The selected clusters are referred to as *EM topoclusters*.

4.1.2 Track reconstruction

Charged particles on traversing the layers of the tracker produce "hits", which are the smallest building blocks of a track. In each layer of the silicon components of the tracker (pixel and SCT), these hits are used to form 3D *space-point* measurements, from which tracks can be reconstructed. Once sets of space points have been formed, the next step is to perform pattern recognition whereby track candidates are identified using a pion hypothesis energy loss model. If this fails, the hypothesis is modified to allow for up to 30% energy loss due to Bremsstrahlung. These track candidates are fitted and in any for which the fit fails with sufficiently small η separation between track and EM cluster, a fit under the electron hypothesis is attempted. This fit allows for more energy loss from Bremsstrahlung. An ambiguity resolution is performed to limit the number of tracks which share space points [30]. Tracks are ranked according to their score, which is calculated using measures of track quality. After ambiguity resolution, track candidates failing basic quality criteria are rejected. These criteria set requirements on the transverse momentum and $|\eta|$ of a track, namely $p_{\mathrm{T}} > 400$ MeV and $|\eta| < 2.5$, as well as the number of space-points and number of space points which are shared between two tracks [30]. The fit for resulting track candidates is then extended to the TRT.

4.1.3 Track-cluster matching

Electron candidates are formed by the matching of tracks with clusters. The fitted tracks are extrapolated from their perigees[1] to the second layer of the EM calorimeter. A track with reconstructed charge q is matched to a cluster where the difference between the azimuthal angles of the track and cluster satisfy $-0.10 < q(\phi_{\mathrm{track}} - \phi_{\mathrm{cluster}}) < 0.05$ [31] and the pseudorapidity difference satisfies $|\eta_{\mathrm{track}} - \eta_{\mathrm{cluster}}| < 0.05$. The asymmetric ϕ requirement is to account for possible Bremsstrahlung radiation by electrons, which can be missed by tracks. If multiple tracks are matched, they are ranked with those having hits in the pixel detector ranking higher than those with hits in only the SCT and TRT. Those with a smaller ΔR, as defined in Equation 3.2, rank higher unless $\Delta R < 0.01$. The highest ranked track is used to reconstruct candidate electrons. If the track has at least 4 hits in the pixel and SCT layers and no vertex from a γ conversion, it is classified as an electron candidate. If the track has no pixel hits and an associated vertex consistent with a conversion in either the SCT or TRT layers, it is classified as

[1]The perigee of a track is defined as the point on a track of closest approach to the interaction point.

a converted photon candidate and if the cluster has no matched track it is classified as an unconverted photon candidate.

4.1.4 Supercluster formation

EM topoclusters from the same electron/photon are collected into a single *supercluster* [27]. This process allows energy from Bremsstrahlung and secondary EM showers to be accounted for. The process works by starting from a seed EM topocluster, to which satellite topoclusters are added. This group of clusters, the seed along with satellites, is referred to as a supercluster. To select seeds, the EM topoclusters are sorted by descending transverse energy, E_{T}, with each cluster being checked individually against criteria. These criteria are that the cluster must not have already been assigned to another supercluster, it must have $E_{\mathrm{T}} > 1$ GeV in the case of electrons or $E_{\mathrm{T}} > 1.5$ GeV in the case of photons and, for electrons, the matched track must have at least 4 silicon hits (pixel or SCT). Topoclusters are added to the supercluster if they are within a 3×5 cell window around the energy weighted barycentre of the seed cluster, which corresponds to a distance of 0.075×0.125 in $\Delta\eta \times \Delta\phi$ space. These topoclusters tend to be caused by secondary EM showers. In the case of electrons, topoclusters are also added if they fall with a 5×12 cell window, corresponding to $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$. This ensures that radiated Bremsstrahlung photons are also added to the supercluster. In the case of converted photons with silicon hits, an electron topocluster is added if the track matched to it shares the same conversion vertex as the seed.

The same track matching procedure as previously is applied to the superclusters, which creates the objects used for analysis.

4.2 Electron-photon calibration

After reconstruction, the energy deposited by an electron or photon is accurately measured. However, this is not necessarily the energy of the original electron/photon. For this reason, the deposited energy must be calibrated to that of the original electron/photon.

4.2.1 Overview of calibration procedure

The calibration of the energy of the original electron/photon in an event begins from the measured energy of the cluster, E_{raw}. Events are simulated using Monte Carlo samples (MC). The calibration procedure is trained on the MC and applied to data,

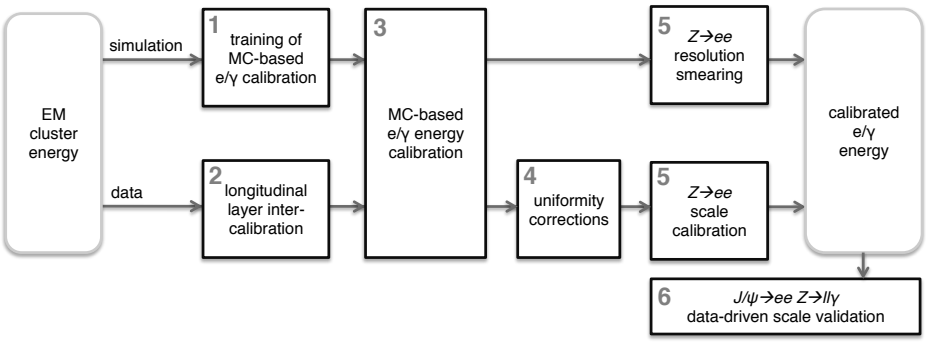


Fig. 4.1 Flowchart showing an overview of the electron and photon calibration procedure. [32]

with some additional corrections applied solely to data or MC. A schematic of the calibration procedure is shown in Figure 4.1. A summary of each step is as follows:

1. **Training of MC-based e/γ calibration**

   The purpose of this step is to train a multivariate analysis (MVA) technique to convert from E_{raw} to the "true" energy of the original electron/photon. The MVA is intended to correct for losses upstream and downstream of the calorimeter and energy deposited in neighbouring cells in the calorimeter which were not included in the cluster [23, 27, 32]. The MVA of choice is a boosted decision tree (BDT) [33], with one tree trained for each of the objects electrons, unconverted photons and converted photons in intervals of η and E_{T}. MC samples of single particle showers with no pile-up are used to train each BDT. The variables used as input are E_{raw}, the energy deposited in the pre-sampler (E_0), the ratio of the energy deposited in the first and second layers of the electromagnetic calorimeter ($E_{1/2}$), the η of the impact point in the calorimeter and the $\eta - \phi$ distance of the closest calorimeter cell from this point. The impact point is defined as the energy-weighted barycentre of the cells in the cluster. Additionally in the case of converted photon candidates, the following variables are also added to the BDT: the estimated radius of conversion in the transverse plane and track properties. Objects for which $1.4 \leq |\eta| < 1.6$ must traverse between 5 - $10X_0$ of material before reaching an active layer of the calorimeter. To account for this, for all objects in the region $1.4 < |\eta| < 1.6$, the energy deposited in the cryostat scintillator (E_4) is also added.

2. **Longitudinal layer intercalibration**

This correction is only applied to data, and is intended to account for the residual difference between data and MC. For this step two calibrations need to be performed. The first is the relative calibration of the first and second layer of the calorimeter and the second is the energy scale of the pre-sampler.

The relative calibration of the first and second layer is performed using muon deposits in the EM calorimeter from $Z \to \mu\mu$ decays. To do this, an $|\eta|$ dependent correction is applied to the energy deposited in the second layer of the EM calorimeter, E_2, of the form $E_2^{\text{corr}} = E_2\alpha_{1/2}$. The correction factor $\alpha_{1/2}$ is given by

$$\alpha_{1/2} = \frac{\left\langle E_{1/2} \right\rangle^{\text{data}}}{\left\langle E_{1/2} \right\rangle^{\text{MC}}}, \tag{4.2}$$

where $\langle E_{1/2} \rangle$ is the ratio of the most probable values (MPVs) of the energy deposited in the first and second layers of the calorimeter [32]. The value of $\alpha_{1/2}$ is calculated using two methods. The first is an analytic fit to the E_1 and E_2 distributions in data and MC using the convolution of a Landau function with a Gaussian. The purpose of the former function is to model the energy deposit and the latter, to model the electronic noise. The second method uses a truncated mean, whereby the mean of each distribution is calculated over a window which excludes the tails. The value of $\alpha_{1/2}$ is taken to be the mean of the value obtained from both approaches, with the difference being the systematic uncertainty.

The pre-sampler energy scale (α_{ps}) is defined as the ratio of the deposited energy in the pre-sampler vs η in data and MC using electrons from decays of W and Z bosons. This must be corrected for mismodelling of passive material, which is applied as

$$\alpha_{\text{ps}} = \frac{E_0^{\text{data}}(\eta)}{E_0^{\text{corr}}(\eta)} = \frac{E_0^{\text{data}}(\eta)}{E_0^{\text{MC}}(\eta)} \times \frac{1}{1 + A(\eta)\left(\frac{E_{1/2}^{\text{data}}(\eta) - 1}{E_{1/2}^{\text{MC}}(\eta) b_{1/2}(\eta)} - 1\right)}, \tag{4.3}$$

where $E_0^{\text{data}}(\eta)$, $E_0^{\text{MC}}(\eta)$ and $E_0^{\text{corr}}(\eta)$ are the measured energies deposited in the pre-sampler vs η for data, MC and corrected MC respectively. The denominator of the rightmost term is the correction for material mismodelling. Here $b_{1/2}(\eta)$ is the ratio of $E_{1/2}$ in data and MC for unconverted photons and $A(\eta)$ is the correlation between E_0 and $E_{1/2}$ when varying the amount of passive material upstream from the pre-sampler.

3. **MC-based e/γ energy calibration**

At this step, the calibration derived in step 1 is applied to the data and MC.

4. **Uniformity corrections**

   Applied only to data, this is to correct for the fact that the detector response is not perfectly uniform vs ϕ in some $\eta - \phi$ regions. There are two effects which cause this. Firstly, in some LAr gaps short circuits can occur, for which reason the high voltage across the gap is set to a non-nominal value. Secondly, some energy is lost between barrel calorimeter modules. Due to the effect of gravity, this is greater for $\phi > 0$. Once these corrections have been applied, a validation is performed by verifying that the invariant mass distribution from $Z \to ee$ events is uniform vs ϕ in the $\eta - \phi$ regions which have been corrected.

5. $Z \to ee$ **resolution smearing and** $Z \to ee$ **scale calibration**

   The two final corrections to be made are corrections to the energy scale for data and corrections to the energy resolution for MC. These are performed separately for different η regions, which will be denoted i. The difference between the energy scale in data and MC in a region i is denoted α_i. The correction to the energy scale is applied as

$$E^{\text{data, corr}} = \frac{E^{\text{data}}}{1 + \alpha_i}, \tag{4.4}$$

   where E^{data} and $E^{\text{data, corr}}$ are the uncorrected and corrected energies from data respectively. In $Z \to ee$ events, the invariant mass from data with one electron in η region i and the second in the region j, m_{ij}^{data}, is corrected as

$$m_{ij}^{\text{data, corr}} = \frac{m_{ij}^{\text{data}}}{1 + \alpha_{ij}}, \tag{4.5}$$

   where α_{ij} is at first order approximated as

$$\alpha_{ij} = \frac{1}{2}\left(\alpha_i + \alpha_j\right). \tag{4.6}$$

   The correction to the resolution is applied as

$$\left(\frac{\sigma_E}{E}\right)^{\text{MC, corr}} = \left(\frac{\sigma_E}{E}\right)^{\text{MC}} \oplus c_i, \tag{4.7}$$

where c_i is the constant energy resolution correction for an η region i and $\oplus$ is a quadrature sum. In a similar manner to before, the mass resolution is therefore

$$\left(\frac{\sigma_m}{m}\right)_{ij}^{\text{MC, corr}} = \left(\frac{\sigma_m}{m}\right)_{ij}^{\text{MC}} \oplus c_{ij}, \tag{4.8}$$

with

$$c_{ij} = \frac{1}{2}\left(c_i \oplus c_j\right). \tag{4.9}$$

Two methods are used to determine the corrections α_{ij} and c_{ij}. The primary method is to optimise the agreement between data and MC by minimising the χ^2 between histograms of the invariant mass distributions of $Z \to ee$ for data and MC in each i, j category. A simultaneous fit is then used to determine the values of α_i and c_i. The second method used is to fit the sum of three Gaussians to the distribution in each category. The differences between the fitted means and widths in data and MC are used to determine each of α_i and c_i. The first method is used to determine the value of the corrections and the difference between the two methods, the systematic uncertainty.

6. $J/\psi \to ee$ $Z \to \ell\ell\gamma$ **data-driven scale validation**

Effects such as the mismodelling of material or shower shapes could cause the calibration to be incorrect. To verify the accuracy of the calibration, two cross checks are used. The first is to perform the entire calibration using electrons from $J/\psi \to ee$ decays. As the J/ψ is much lighter than the Z, the electrons resulting from the decay will be much lower energy, allowing an independent cross check. From this procedure, values of residual energy scale differences, $\Delta\alpha_i$, are computed. These are defined as the difference between the two values of α_i derived by each method. If these are consistent with $\Delta\alpha_i = 0$ for all i, the calibration is correct.

The second cross check verifies that the calibration is also valid for photons. The energy scale calibration is performed again, using $Z \to \ell\ell\gamma$ events and residual energy scale differences are calculated. This is done separately for each of $\ell = e, \mu$ and converted and unconverted photons. As before, if $\Delta\alpha_i = 0$ for all i, the calibration is correct.

4.2.2 Further checks on the photon calibration procedure

Accurate calibration of photons is necessary for measurements of the Higgs boson properties in the $H \to \gamma\gamma$ channel, as the final state is composed entirely of photons, and in the $H \to ZZ^* \to 4\ell$ channel, as final state radiation photons can be emitted. In the case of $H \to \gamma\gamma$ measurements, accurate calibration of photons is particularly important as measurements are systematic uncertainty dominated [6]. The impact of the systematic uncertainty arising from the photon energy scale calibration on the measurement of m_{H} was 0.21-0.36% in the early Run 2 ATLAS measurement [34].

For electrons, energy scale calibration is performed using the $Z \to ee$ resonance as this peak has little background and the mass of the Z boson is precisely known. No such diphoton resonance exists, so the scale corrections for electrons and photons are assumed to be the same. As mentioned above, ATLAS uses $Z \to \ell\ell\gamma$ decays to validate the photon calibration. The development of a further check on the photon calibration procedure is presented here. Note that the number of events used in samples used in this study is relatively low as the work is intended as a proof-of-concept.

To perform this check, electrons from $Z \to ee$ events are "transformed" into unconverted photons by calibrating them as such and the consistency between data and MC as well as real and transformed photons is verified. These checks are performed using data and MC $Z \to ee$ samples from 13 TeV pp collisions corresponding to an integrated luminosity of 81 fb^{-1}. The following selections are then applied to these samples.

- Single lepton triggers
- Medium identification requirements (for an explanation of identification requirements in the case of electrons , see section 4.3)
- $E_{\mathrm{T}} > 25$ GeV
- Exactly one topocluster

Next the data and MC electrons are recalibrated as unconverted photons. These photons from transformed electrons will be referred to as *transformed photons* or γ^{TR}. The data/MC difference is compared using the variable $(\langle E_e/E_{\mathrm{raw}}\rangle^{\mathrm{MC}} - \langle E_e/E_{\mathrm{raw}}\rangle^{\mathrm{data}}) - (\langle E_{\gamma^{\mathrm{TR}}}/E_{\mathrm{raw}}\rangle^{\mathrm{MC}} - \langle E_{\gamma^{\mathrm{TR}}}/E_{\mathrm{raw}}\rangle^{\mathrm{data}})$, where $\langle E_e/E_{\mathrm{raw}}\rangle^{\mathrm{MC}}$ is the average value of the ratio of the calibrated electron energy to the raw energy reconstructed in the EM calorimeter for MC, $\langle E_e/E_{\mathrm{raw}}\rangle^{\mathrm{data}}$ is that of data and $\langle E_{\gamma^{\mathrm{TR}}}/E_{\mathrm{raw}}\rangle^{\mathrm{MC}}$ and $\langle E_{\gamma^{\mathrm{TR}}}/E_{\mathrm{raw}}\rangle^{\mathrm{data}}$ are the same ratios for transformed photons for MC and data respectively. Distributions of

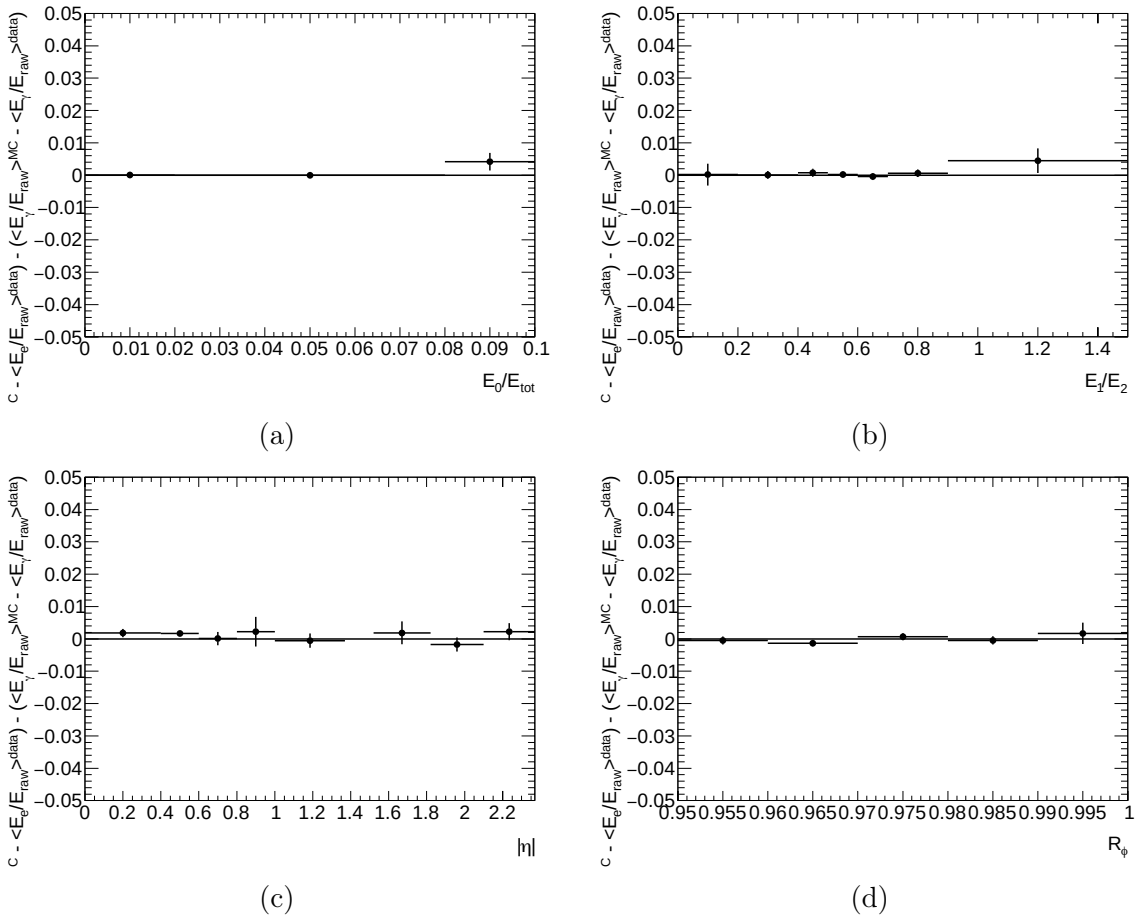
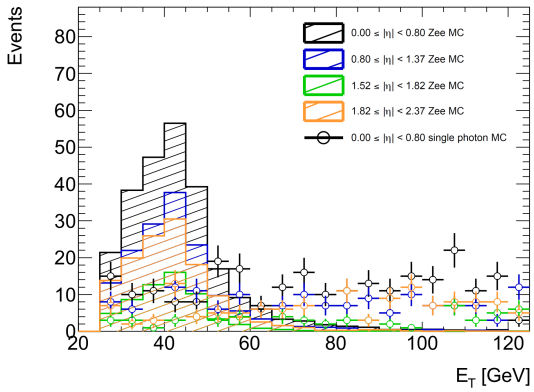


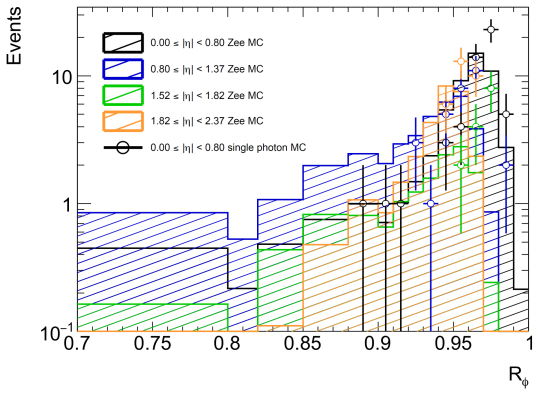
Fig. 4.2 $(\langle E_e/E_{\rm raw}\rangle^{\rm MC} - \langle E_e/E_{\rm raw}\rangle^{\rm data}) - (\langle E_{\gamma^{\rm TR}}/E_{\rm raw}\rangle^{\rm MC} - \langle E_{\gamma^{\rm TR}}/E_{\rm raw}\rangle^{\rm data})$ vs (a) fraction recovered energy in the EM calorimeter deposited in the pre-sampler, $E_0/E_{\rm tot}$, (b) ratio of energy deposited in the first layer of the EM calorimeter to the second layer, E_1/E_2, (c) $|\eta|$ and (d) ratio of the energy deposited in the 3×3 cells centred at the cluster to the deposit in 3×7, R_ϕ. The choice of y-axis scale here is to allow the same scale to be used as in Figure 4.4.

$(\langle E_e/E_{\rm raw}\rangle^{\rm MC} - \langle E_e/E_{\rm raw}\rangle^{\rm data}) - (\langle E_{\gamma^{\rm TR}}/E_{\rm raw}\rangle^{\rm MC} - \langle E_{\gamma^{\rm TR}}/E_{\rm raw}\rangle^{\rm data})$ vs shower shape variables are shown in Figure 4.2. Here, the agreement between data and MC is good, so the calibration procedure is consistent between photons and electrons.

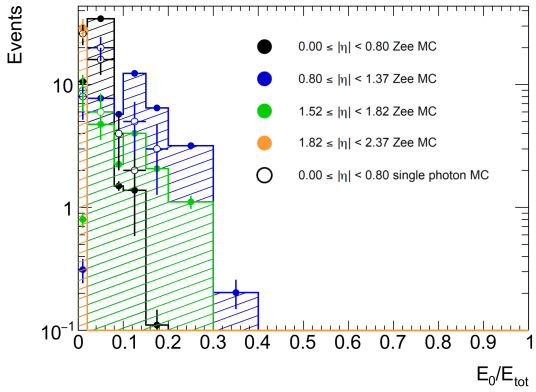
The next check is to verify the consistency between real photons and transformed photons. For this test, only the most photon-like electrons from $Z \to ee$ MC events are selected. To select these electrons, they are compared to photons from single photon MC. The selection used here is exactly one unconverted photon with $E_{\rm T} > 25$ GeV satisfying strict identification requirements in each event. The same $Z \to ee$ selections as in the previous test are used. Before transforming the photons, additional selections are applied to the electrons from $Z \to ee$ and the real photons using kinematic and



(a) E_{T}



(b) R_{ϕ}

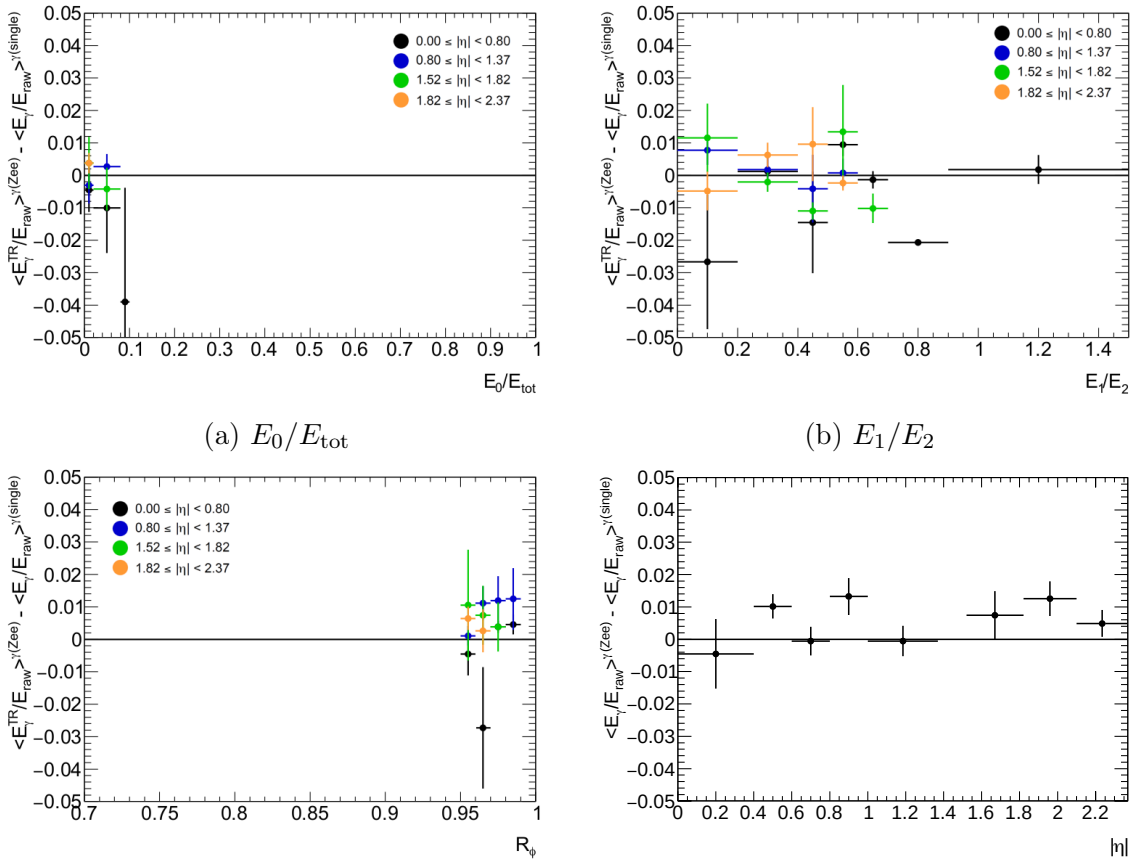


(c) E_0/E_{tot}

Fig. 4.3 Distributions of kinematic and shower shape variables for electrons from $Z \to ee$ MC and photons from single photon MC. All figures show $Z \to ee$ with selections of single lepton triggers, exactly two electrons satisfying medium identification requirements, with both having $E_{\mathrm{T}} > 25$ GeV and exactly one topocluster and photon selections of a single unconverted photon with $E_{\mathrm{T}} > 25$ GeV satisfying strict identification criteria. Figures (b) and (c) have the additional requirement of $E_{\mathrm{T}} < 50$ GeV for both electrons and photons. In each figure, the filled region is the distribution for electrons and the empty circles are that of the photons.

shower shape variables. Figure 4.3a shows the E_{T} distributions of the $Z \to ee$ electrons and real photons. The electrons tend to have $E_{\mathrm{T}} < 50$ GeV whereas photons have no dependence. As photons outside the energy range of the $Z \to ee$ electrons are not of interest, photons and electrons with $E_{\mathrm{T}} < 50$ GeV are selected. This selection is applied, the distributions of shower spread in the ϕ direction (R_{ϕ}) and the ratio of the energy deposited in the pre-sampler (E_0/E_{tot}) of the photons and electrons are compared as shown in Figures 4.3b and 4.3c respectively. Figure 4.3b shows that electromagnetic showers from photons tend to have a smaller spread in the ϕ direction

than electrons, so an additional selection is applied to the electrons requiring that $R_\phi \geq 0.95$. Figure 4.3c shows that photons tend to deposit a smaller fraction of their energy in the pre-sampler than electrons. So, electrons with $E_0/E_{\text{tot}} < 0.1$ are selected. Now that a selection of photon-like electrons has been obtained, the remaining electrons from $Z \to ee$ are transformed to photons and the energy correction is verified. To verify



(a) E_0/E_{tot} (b) E_1/E_2

(c) R_ϕ (d) $|\eta|$

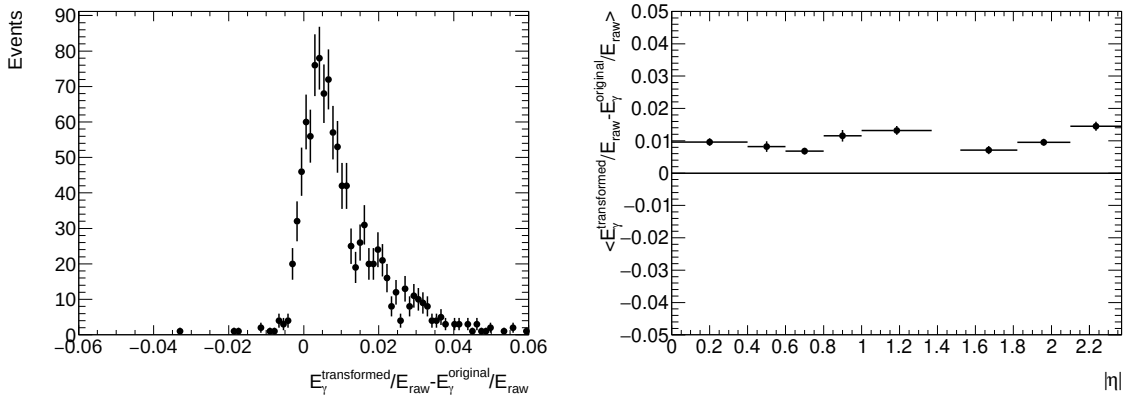
Fig. 4.4 The difference the mean ratio of calibrated to raw energy between transformed and real photons, $\langle E_{\gamma^{\text{TR}}}/E_{\text{raw}}\rangle^{Z\to ee} - \langle E_\gamma/E_{\text{raw}}\rangle^{\text{single}-\gamma}$, vs shower shape variables. The variation for each variable is shown in four bins of $|\eta|$.

this correction, the difference in the mean ratio of calibrated to raw energy between transformed and real photons, $\langle E_{\gamma^{\text{TR}}}/E_{\text{raw}}\rangle^{Z\to ee} - \langle E_\gamma/E_{\text{raw}}\rangle^{\text{single}-\gamma}$, is investigated vs shower shape variables. Figure 4.4 shows the results of this. These are consistent with the expected value of 0, so the calibration procedure is working as expected.

4.2.3 Investigating the effect of misidentifying photon conversions

Each unconverted (converted) photon reconstructed by ATLAS has a chance to be misidentified as converted (unconverted). The following is an investigation into the effect on the calibrated photon energy of misidentifying unconverted photons as photons which convert in the TRT and vice versa. There are two possibilities for a photon that converts in the TRT, that is it will leave either one or two tracks. In the latter case, both the electron and positron produced in the conversion leave a track in the TRT and leave clusters in the calorimeter. These are referred to as *double TRT converted* photons. In the case of those leaving one track, either the electron or positron produced may be of very low transverse momentum and will therefore not be detected. These are referred to as *single TRT converted* photons. This misidentification can lead to a bias in the energy calibration as tracks and clusters which do not belong to a particular photon can be added to the photon object.

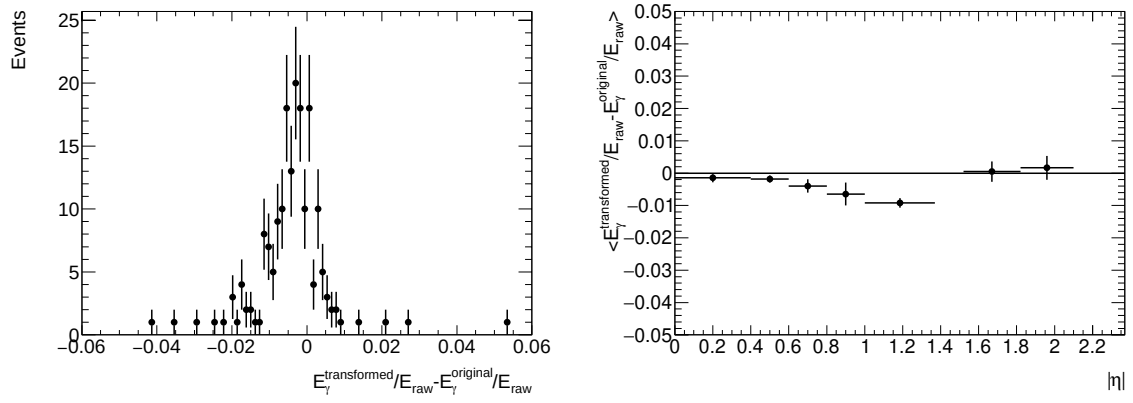
First the effect of misidentifying unconverted photons as single or double TRT converted photons is investigated. For this, single photon MC is used, selecting events with exactly one unconverted photon with $E_{\mathrm{T}} > 25$ GeV satisfying strict identification criteria. Half of the remaining photons are transformed into single TRT converted



(a) (b)

Fig. 4.5 Behaviour of $E_\gamma^{\mathrm{transformed}}/E_{\mathrm{raw}} - E_\gamma^{\mathrm{original}}/E_{\mathrm{raw}}$, where E_{raw} is the uncalibrated (raw) energy of the photon, E_{original} refers to the calibrated energy of MC unconverted photons and $E_\gamma^{\mathrm{transformed}}$ to that of the corresponding photon transformed to either a single or double TRT converted photon with 50% probability. (a) Shows the distribution and (b) shows the behaviour vs $|\eta|$.

photons and the other half to double TRT converted. To determine the extent of the miscalibration this would cause, the difference between the ratios of the calibrated to raw



(a) (b)

Fig. 4.6 Behaviour of $E_\gamma^{\text{transformed}}/E_{\text{raw}} - E_\gamma^{\text{original}}/E_{\text{raw}}$, where E_{raw} is the uncalibrated (raw) energy of the photon, E_{original} refers to the calibrated energy of the MC single/double TRT converted photon and $E_\gamma^{\text{transformed}}$ to that of the corresponding photon transformed to an unconverted photon. (a) Shows the distribution and (b) shows the behaviour vs $|\eta|$.

energies for the transformed and original photon sets, $E_\gamma^{\text{transformed}}/E_{\text{raw}} - E_\gamma^{\text{original}}/E_{\text{raw}}$ are determined. The distribution of this is shown in Figure 4.5a and the behaviour vs $|\eta|$ in Figure 4.5b. These show that incorrectly reconstructing all unconverted photons as single or double TRT converted would result in a bias in energy calibration of $\sim +1\%$, independent of η. That is, an unconverted photon misidentified as converted will on average have a calibrated energy $\sim 1\%$ larger than its "true" energy. It should be noted the the effect on the ATLAS calibration is expected to be much smaller than this figure. To quantify the impact of such a misidentification on the ATLAS photon calibration, the misidentification rate of unconverted photons as single/double TRT converted photons would need to be calculated.

A similar test is also performed to check the bias resulting from incorrectly reconstructing single and double TRT converted photons as unconverted photons. As previously, a single photon MC sample is used with the requirement that the photons pass strict identification requirements and have $E_{\text{T}} > 25$ GeV. The remaining photons must be either single or double TRT converted and all of these are transformed to unconverted.

The distribution of $E_\gamma^{\text{transformed}}/E_{\text{raw}} - E_\gamma^{\text{original}}/E_{\text{raw}}$, where $E_\gamma^{\text{original}}$ is now the calibrated energy of the initially single/double converted TRT photon and $E_\gamma^{\text{transformed}}$ is the calibrated energy of the transformed unconverted photon, is shown in Figure 4.6a. The bias in energy calibration from performing this transformation is $< 1\%$.

$E_\gamma^{\text{transformed}}/E_{\text{raw}} - E_\gamma^{\text{original}}/E_{\text{raw}}$ vs $|\eta|$ is shown in Figure 4.6b. This shows that bias is η dependent, with the largest bias of $\sim -1\%$ occurring in the region $0.8 \leq |\eta| < 1.37$.

4.3 Electron identification

The purpose of electron identification is to reduce contamination in data samples of electrons from secondary decays, referred to as *non-prompt* electrons, and other objects misidentified as electrons. The strictness of these requirements can be chosen depending on the requirements of an analysis. This is done by providing a series of *operating points* which relate to sets of selections. These are, in order of increasing purity and rejection of "true" electrons, referred to as *very loose*, *loose*, *medium* and *tight*.

4.3.1 Likelihood discriminant

Accurate electron identification requires a good separation between *prompt* signal electrons with those originating from background processes. These background non-prompt electrons include misidentified hadrons, electrons from photon conversions and non-isolated electrons from decays of heavy flavour hadrons.

Electrons are separated from background using a likelihood-based discriminant. The likelihood for signal (background) is given by

$$L_{S(B)}(x_i) = \prod_{i=1}^{n} P_{S(B),i}(x_i), \tag{4.10}$$

where $P_{S(B),i}(x_i)$ is value of the the signal (background) PDF for the variable i of value x_i [31]. A list of variables used can be found in Table 4.1. Correlations between variables are neglected. For each electron candidate, a discriminant is defined as

$$d_L = \frac{L_S}{L_S + L_B}, \tag{4.11}$$

where L_S and L_B are the likelihoods for the signal and background respectively, as given by Equation 4.10. This discriminant gives a sharp peak at 1 for signal and 0 for background. As operating points need to be determined precisely, such behaviour is undesirable as it would require fine binning of the discriminant distribution in order to achieve this. To broaden the distribution of d_L, it is transformed by an inverse sigmoid

Name	Description	Correction	
		Shift	Width
d_0	Transverse impact parameter, the distance from the track to the beam spot at closest approach.		
$d_0^{\text{significance}}$	Significance of the transverse impact parameter $= \lvert d_0/\sigma(d_0)\rvert$, where $\sigma(d_0)$ is the uncertainty on d_0.	✓	
$\Delta\eta_1$	$\Delta\eta$ between cluster in first layer of the EM calorimeter and extrapolated track.	✓	✓
$\Delta\phi_{\text{res}}$	$\Delta\phi$ between the cluster in the second layer of the EM calorimeter and the momentum-rescaled track, extrapolated from the perigee, multiplied by the charge.	✓	✓
$\Delta p/p$	Momentum lost by the track between the perigee and the last measurement point divided by the momentum at perigee.	✓	
E_{ratio}	The difference of the energy deposited in the maximum deposit and that of a secondary maximum divided by the sum of the two energies.	✓	
f_1	Fraction of energy deposited in the first layer of the EM calorimeter to total energy deposited in the EM calorimeter.	✓	✓
f_3	Fraction of energy deposited in the third layer of the EM calorimeter to total energy deposited in the EM calorimeter.	✓	✓
R_η	Ratio of the energy deposited in the 3×7 cells centred at the cluster to the deposit in 7×7.	✓	
R_ϕ	Ratio of the energy deposited in the 3×3 cells centred at the cluster to the deposit in 3×7.	✓	
R_{had}	Hadronic leakage: ratio of E_{T} deposited in the hadronic calorimeter to E_{T} of the EM cluster. Used for electrons outside the calorimeter transition region.	✓	✓
R_{had1}	Hadronic leakage: ratio of E_{T} deposited in the first later of the hadronic calorimeter to E_{T} of the EM cluster. Used only for electrons inside the calorimeter transition region.		
$w_{\eta 2}$	Lateral shower width: $\sqrt{(\sum E_i\eta_i^2)/(\sum E_i) - ((\sum E_i\eta_i)/\sum E_i)^2}$, where E_i and η_i are the energy deposited and pseudorapidity of cell i, taken only in a 3×5 cell window, centred on the cluster.	✓	

Table 4.1 Variables used in the training of the likelihood discriminant. Variables defined in e.g. "3×3 cells" refer to the cells of the second layer of the electromagnetic calorimeter in $\Delta\eta \times \Delta\phi$ space in steps of 0.025. The two rightmost columns indicate whether shift or width corrections are applied to the PDF of the variables to correct for shower shape mis-modelling.

function as

$$d'_L = -\tau^{-1} \ln(d_L^{-1} - 1), \tag{4.12}$$

where τ takes a value of 15 [35].

To train the likelihood, accurate signal-background separation is required over a wide E_{T} range. For this training, electrons from $J/\psi \to ee$ decays are used for $E_{\mathrm{T}} < 15$ GeV and from $Z \to ee$ decays for $E_{\mathrm{T}} \geq 15$ GeV. Backgrounds considered are multijet production, top quark production and electroweak processes. All of these are simulated using MC.

To study electrons in an event, a *tag and probe* method is used whereby one of the electrons, the tag, is required to satisfy strict selections and the second, the *probe*, is studied. The tight selections on the tag allow a low-background sample to be selected and the loose selections on the probe ensure that the studied sample of electrons is unbiased. $J/\psi \to ee$ events are required to have two reconstructed electron candidates with $E_{\mathrm{T}} > 4.5$ GeV falling within $|\eta| < 2.47$ with the event passing dielectron triggers and having a dielectron invariant mass between 1 and 6 GeV. At least one of the electron candidates (the tag) must pass the tight likelihood threshold and not fall in the calorimeter transition region $1.37 < |\eta| < 1.52$. The remaining electron candidate serves as the probe. In the case that both electrons satisfy the strict tag requirements, both electrons serve as probes. $Z \to ee$ events are required to have two reconstructed electron candidates with $|\eta| < 2.47$ with at least one of these satisfying single electron triggers. The invariant mass of the two electron candidates must be within 15 GeV of the Z boson mass. The tag electron must be the electron candidate which satisfied the triggers, not fall within the calorimeter transition region $1.37 < |\eta| < 1.52$ and have $E_{\mathrm{T}} > 27$ GeV. In the case that both satisfy the strict tag requirements, both electrons serve as probes.

E_{T} bin boundaries [GeV]												
$J/\psi \to ee$				$Z \to ee$								
4.5	7	10	15	15	20	25	30	35	40	45	80	150	∞

Table 4.2 Bin boundaries in E_{T} used for electron identification. For the lower E_{T} bins ($E_{\mathrm{T}} \leq 15$ GeV), $J/\psi \to ee$ events are used and $Z \to ee$ for the higher E_{T} (> 15 GeV).

$\lvert\eta\rvert$ bin boundaries									
0.00	0.60	0.80	1.15	1.37	1.52	1.81	2.01	2.37	2.47

Table 4.3 Bin boundaries in $|\eta|$ used for electron identification.

4.3.2 Shower shape corrections

The shapes of electromagnetic showers are poorly modelled in ATLAS, the reason for which is presently unknown. The effect of this mis-modelling is that some of the distributions of shower shape variables differ between MC simulations and data. To correct for this, the MC distributions of signal shower shape variables are corrected to match the data by the transforming of their histograms. To account for the possibility of the mis-modelling not being uniform across E_{T} and η, the correction is derived separately in bins of E_{T} and $|\eta|$. The binning used is shown in Tables 4.2 and 4.3. Before any correction is applied, the distributions of each variable are smoothed using a method of kernel density estimation (KDE) in order to remove fluctuations [36]. Two

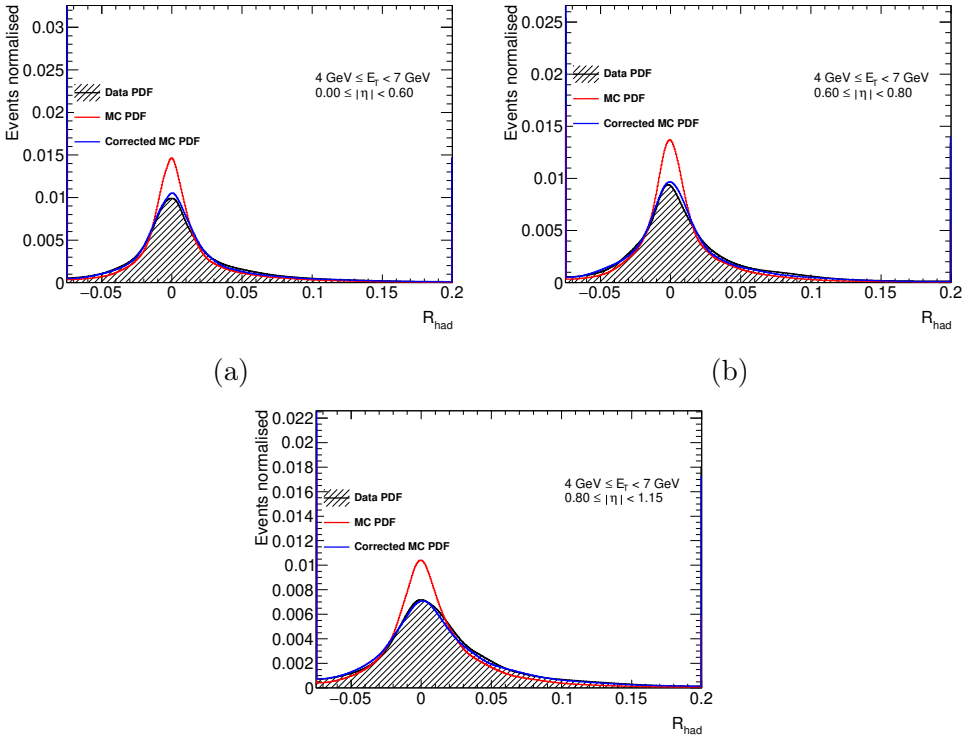


Fig. 4.7 Corrections applied to the variable R_{had} for $4 \leq E_{\mathrm{T}} < 7$ GeV the bin (a) $0.00 \leq |\eta| < 0.60$, (b) $0.60 \leq |\eta| < 0.80$ and (c) $0.80 \leq |\eta| < 1.15$. The data and MC used correspond to $J/\psi \to ee$ decay events. This variable requires a relatively large width correction and a small shift.

corrections can be applied to each smoothed histogram. The first is a width correction, which is done by calculating the full width at half maximum (FWHM) of the simulation and data. The ratio of these is taken and the MC histogram is smeared by this fraction.

In the case where $FWHM_{\mathrm{MC}} > FWHM_{\mathrm{data}}$, no correction is applied. The second

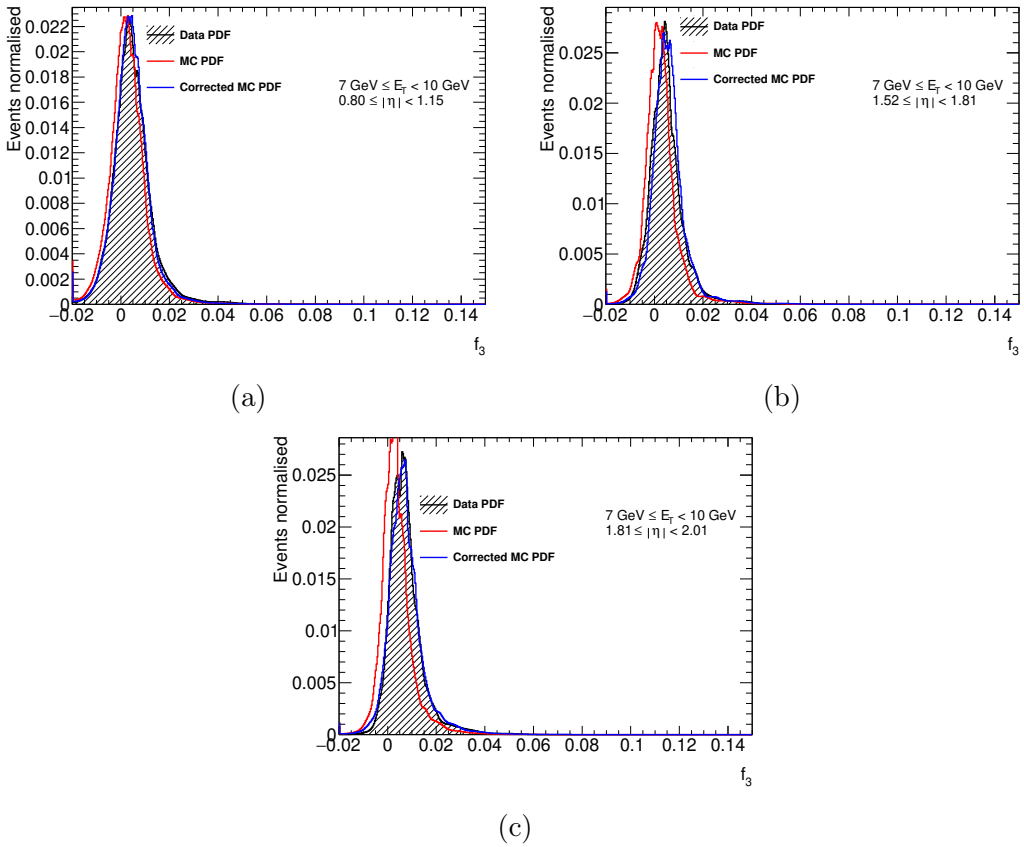


Fig. 4.8 Corrections applied to the variable f_3 for $7 \leq E_{\mathrm{T}} < 10$ GeV left to right in the bin (a) $0.80 \leq |\eta| < 1.15$, (b) $1.52 \leq |\eta| < 1.81$ and (c) $1.81 \leq |\eta| < 2.01$. The data and MC used correspond to $J/\psi \to ee$ decay events. This variable requires a minor shift and width correction.

correction is a shift. This is done by shifting the MC histogram bin-by-bin and using a χ^2 fit to find the optimal shift. Table 4.1 shows which variables have a shift and width correction applied. The Figures 4.7-4.9 demonstrate the effect of corrections in three different cases. The first of these cases, Figure 4.7, shows the corrections to the R_{had} variable in three bins for $4 \leq E_{\mathrm{T}} < 7$, $J/\psi \to ee$. R_{had} requires a large width correction, with the correction factor being between 1.4 and 1.5 in the cases shown. The corrected distribution matches the data very closely. The second case shown in Figure 4.8 is the correction in three $7 \leq E_{\mathrm{T}} < 10$ bins for the variable f_3 for $J/\psi \to ee$. This variable requires a larger shift correction than R_{had}, but a much smaller width correction. As before, post correction, the MC matches the data well. The final case shown in Figure 4.9 is also for the variable f_3, for $Z \to ee$ in three $30 \leq E_{\mathrm{T}} < 40$ bins. As with the previous case, a small width correction and a shift correction are required.

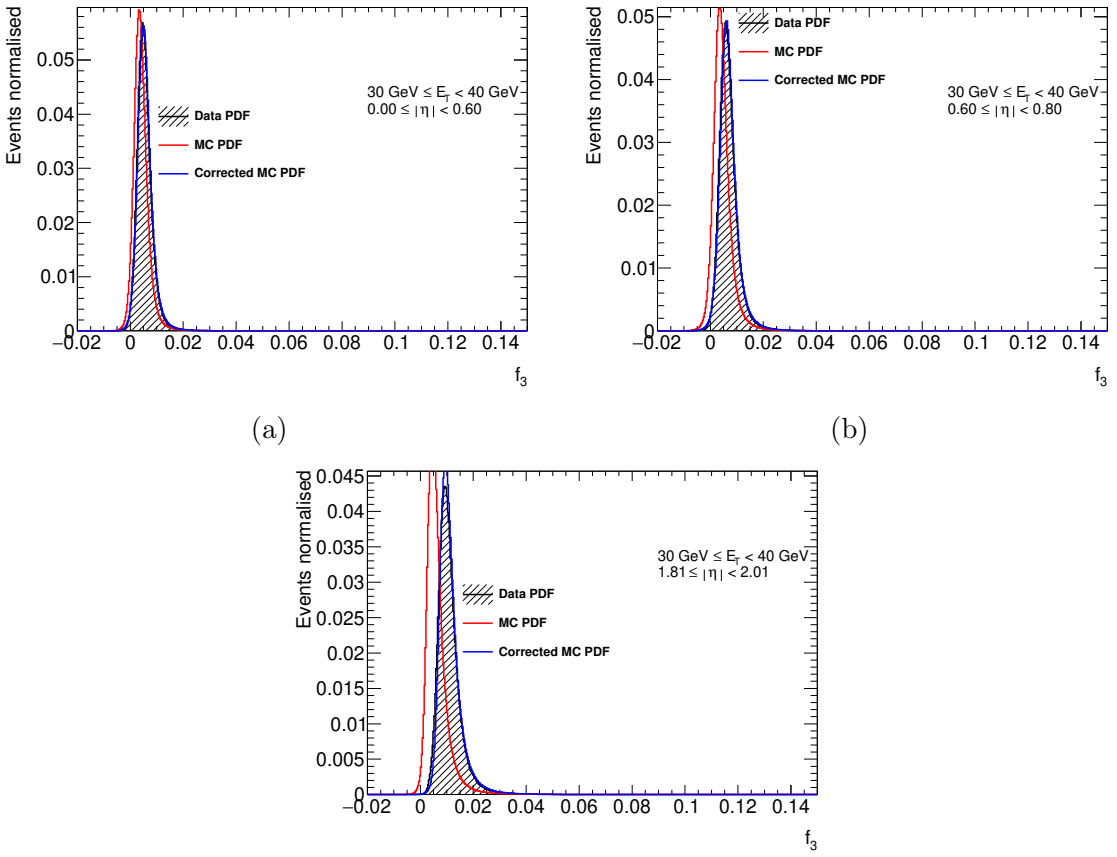


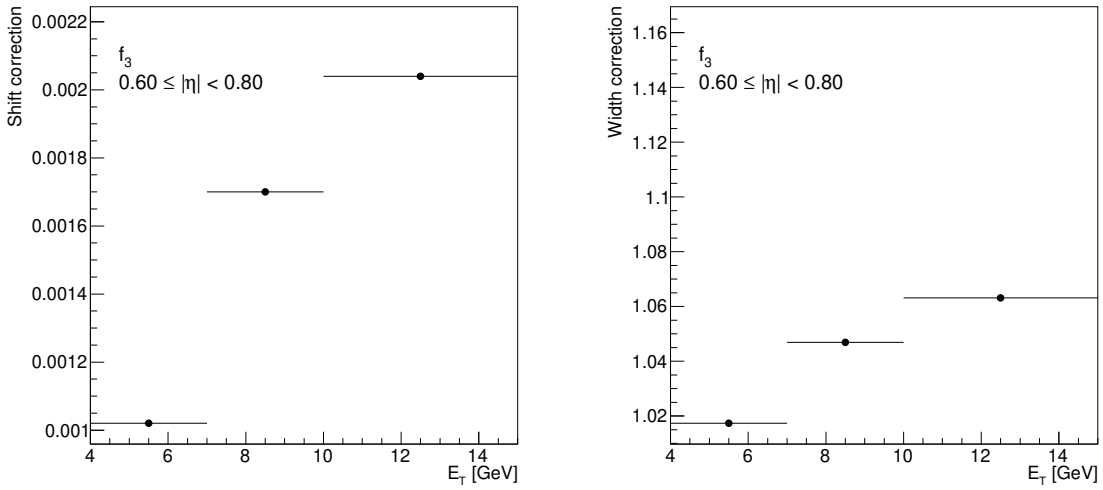
Fig. 4.9 Corrections applied to the variable f_3 for $30 \leq E_{\mathrm{T}} < 40$ GeV in the bin (a) $0.00 \leq |\eta| < 0.60$, (b) $0.60 \leq |\eta| < 0.80$ and (c) $1.81 \leq |\eta| < 2.01$. The data and MC used are from $Z \to ee$ decays. This variable requires a minor shift and width correction.

The shift here in each case is ∼0.002 with width corrections of < 1.13. As in the previous cases, the corrected MC provides a good description of the data. Figure 4.10 shows the variation in correction factor for the variable f_3 vs $|\eta|$ and E_{T}. This variable shows a dependence on both $|\eta|$ and E_{T} for the optimal correction.

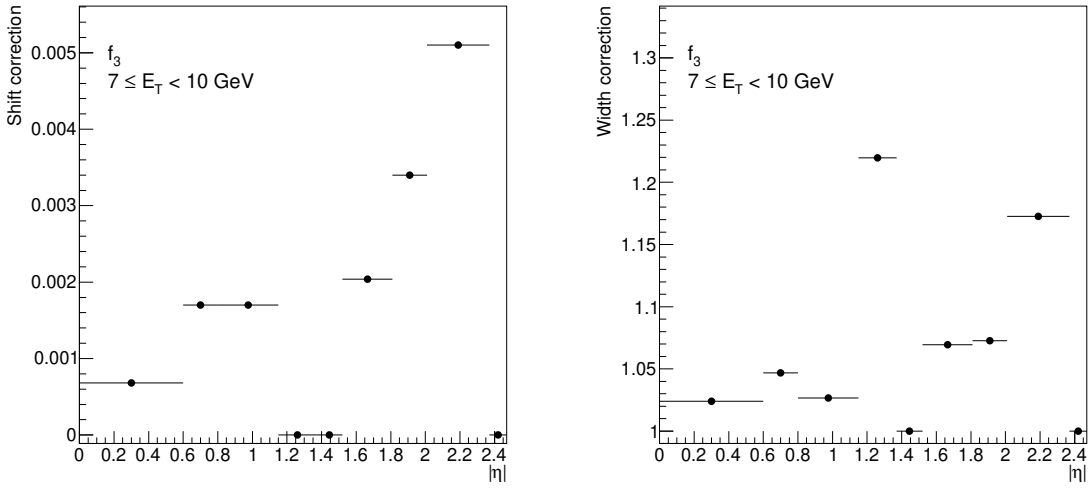
To define the operating points, the desired performance is expressed in terms of target efficiencies. In the tuning of these operating points, *electron ID efficiency* is used, defined as the proportion of reconstructed electron candidates which pass the ID requirements,

$$\varepsilon_{\mathrm{id}} = \frac{N_{\mathrm{id}}}{N_{\mathrm{reco}}}, \tag{4.13}$$

where N_{reco} is the number of electron candidates which are reconstructed and N_{id} is the number which are both reconstructed and identified.



(a) (b)



(c) (d)

Fig. 4.10 Variation in corrections vs $|\eta|$ and E_{T} for f_3 from $J/\psi \to ee$ decays. Shown in (a) and (b) are the variation in shift and width correction vs E_{T} for $0.60 \leq |\eta| < 0.80$ respectively and (c) and (d) shift and width correction vs $|\eta|$ for $7 \leq E_{\mathrm{T}} < 10$ GeV respectively.

Each operating point is tuned by choosing a selection on d'_L which meets a particular pre-defined target efficiency. This is performed separately in bins of η and E_{T}, with each bin having a different target efficiency. In addition to a selection on d'_L, for the Loose, Medium and Tight operating points, the track of the candidate electron must have at least two hits in the pixel layer with at least 7 total in the pixel and SCT. Additionally, Medium and Tight also require that one of the hits be in the innermost pixel layer.

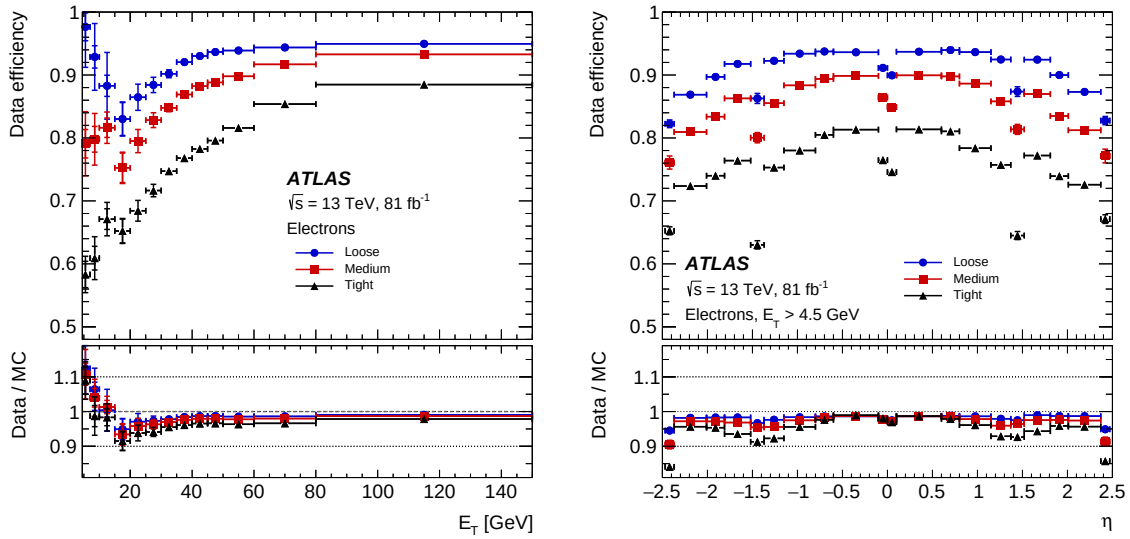


Fig. 4.11 Electron ID efficiency ($\varepsilon_{\mathrm{id}}$) vs E_{T} (left) and η (right) for the Loose, Medium and Tight operating points [27]. The comparison with expected efficiencies from MC are also shown.

The resulting distributions of $\varepsilon_{\mathrm{id}}$ vs E_{T} and η for the Loose, Medium and Tight operating points are shown in Figure 4.11. $\varepsilon_{\mathrm{id}}$ increases smoothly with increasing E_{T}, with the exception of the low E_{T} regions. This is caused by the correction method working better at higher E_{T}, leading to the efficiency in this region being larger than intended.

4.3.3 Improvements to shower-shape corrections

One of the limitations with the method of using FWHM to calculate the width adjustment for histogram is in the case of noisy or bimodal distributions. Here, the calculation of the FWHM is more difficult and can lead to cases of overcorrection. Another case causing overcorrection, is that in some cases, a width and shift correction

is not sufficient to fully correct for the mismodelling. Presented here is a method developed to solve these problems.

As the probability distribution of the sum of two or more independent random variables is the convolution of their individual distributions, the shift and width adjustments described previously can be viewed as a convolution with a Gaussian. So the probability density function of the corrected MC distribution becomes

$$MC_{\text{corrected}}(x) = \int MC(t)G(\mu, \sigma; x - t)dt, \tag{4.14}$$

where $MC(t)$ is the uncorrected MC PDF and $G(\mu, \sigma; x - t)$ is a Gaussian with a mean of μ and width of σ. The parameters of this Gaussian control the adjustments, i.e. μ controls the shift and σ controls the width. For the case where no width adjustment is required, $\sigma \to 0$ so $G(x - t) \to \delta(x - t)$ and in the case where no shift is required, μ is fixed at 0. As previously, the $MC(t)$ and the data are taken in the form of smoothed histograms. The convolution is fitted to the data using a binned maximum likelihood fit with the μ and σ parameters of the Gaussian floating. Width corrections are applied if the RMS of the data is greater than 1.01 times that of the MC and shift corrections are applied if the difference between the maxima of MC and data exceeds the bin width. To reduce the effect of the statistical fluctuations in the tails when correcting each variable, a threshold number of events is set at both the low and high ends of the distribution to define the window in which the correction is performed. Two examples of the convolution method showing better performance than the shift and FWHM method are shown in Figures 4.12 and 4.13. In both of these cases, the shifting and widening of the variables is insufficient to fully correct the MC, so the shift and FWHM method tends to over-correct the width of the distribution. The convolution method however tends not to over-correct, and the correction here serves as a better description of the data.

This method, whilst not yet used in any published ATLAS results, is being used to correct for mismodelling of the calorimeter energy distribution in the isolation of photons.

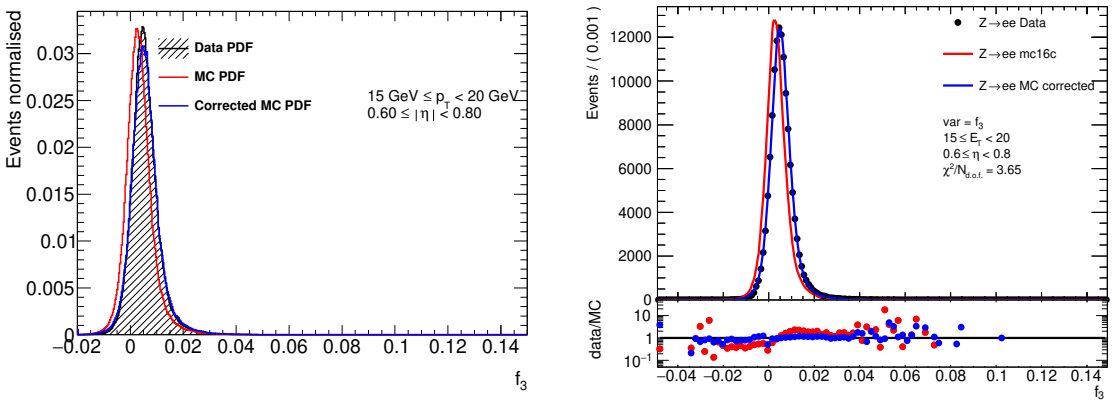


Fig. 4.12 Comparison of the two correction methods shown for the distribution of f_3 in the $15 \leq E_{\mathrm{T}} < 20$, $0.60 \leq |\eta| < 0.80$ bin. These are the shift and FWHM method (left) and the convolution method (right).

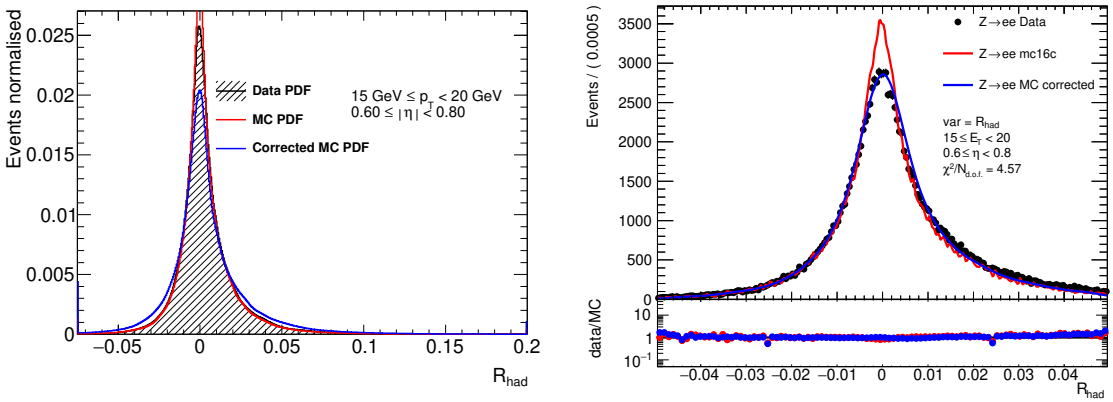


Fig. 4.13 Comparison of the two correction methods shown for the distribution of R_{had} in the $15 \leq E_{\mathrm{T}} < 20$, $0.60 \leq |\eta| < 0.80$ bin. These are the shift and FWHM method (left) and the convolution method (right).

Chapter 5

Standard Model Higgs boson mass measurement in $H \to 4\ell$ at 36.1 fb^{-1}

5.1 Overview

This chapter describes the mass measurement of the Higgs boson in the $H \to ZZ^* \to 4\ell$ channel ($\ell = \mu, e$) using ATLAS data corresponding to an integrated luminosity of 36.1 fb^{-1} from pp collisions at a 13 TeV centre of mass energy collected during LHC early Run 2.

Whilst the mass of the Higgs boson is a free parameter in the standard model, other parameters such as the Higgs boson production cross section and branching ratios depend on this value. So a measurement of the mass along with other properties of the Higgs boson can serve as a test of the standard model.

The strategy is to select events which contain two pairs of same-flavour, opposite-sign electrons (e) or muons (μ) and build a model of the four-lepton invariant mass distribution ($m_{4\ell}$) which can be compared to the data to measure the mass of the Higgs boson (m_{H}). As the decay products are electrons and muons, four different final states are possible, namely four muons (4μ), four electrons ($4e$), two muons and two electrons ($2\mu 2e$) and two electrons and two muons ($2e2\mu$), with the difference between these last two being that in the former case, the Z producing the two muons is on shell and that producing the two electrons is off-shell and vice versa in the case of the latter. Each of these final states will be referred to as a *channel*. The second categorisation is done with use of a BDT designed to distinguish between signal and background events using the kinematics of the four-lepton system. Each of the defined categories is treated separately, with a signal and background model being built independently in each. The resolution of $m_{4\ell}$ is improved using two methods. The first is to account

for final state radiation (FSR) and the second comes from considering the fact that one of the Z bosons in the decay is usually produced on-shell, allowing its mass to be constrained to the true Z lineshape.

The background is described using a *template* method. That is, simulated distributions are used and smoothed to create a smooth non-parametric probability density function (PDF) to describe the shape of the background.

The signal model is created using MC simulations of the $m_{4\ell}$ signal distribution at several discrete values of m_{H} from which, a continuous model of the $m_{4\ell}$ distribution can be built vs m_{H}. In the previous Run 1 analysis, the signal model was built in a similar manner to the background model [37]: that is, smoothed templates were created at several discrete values of m_{H}, which were interpolated to create a continuous picture of $m_{4\ell}$ vs m_{H}. In this analysis, the template signal model is replaced with a *per-event response* model, which builds a description of $m_{4\ell}$ by considering the energy and momentum responses of the leptons in an event. One key advantage of such a model is that it accounts for the per-event uncertainties on an individual measurement of $m_{4\ell}$.

5.2 Simulation of events

In order to build a model of the signal $m_{4\ell}$ lineshape vs m_{H}, the decay of Higgs bosons to four leptons is simulated using Monte Carlo (MC) at various different, discrete values of m_{H} from different production modes. These MC samples generated for different values of m_{H} are referred to as *mass points*.

The production modes considered are gluon-gluon fusion (ggF), vector boson fusion (VBF), vector boson associated production (VH) and $t\bar{t}H$ production. Each of these is discussed in Section 2.7. MC samples for ggF and VBF are used to create the signal model, generated at mass points between 120 and 130 GeV. MC samples for other production modes are calculated only for the 125 GeV mass point and are used only for validation purposes.

The signal yield is estimated by considering the Higgs production cross section, the branching ratio for four-lepton decay and detector acceptance.

The backgrounds to the mass measurement can be divided into two categories, namely *irreducible* and *reducible*. The former is the more important and refers to events with four prompt leptons which pass event selection but do not originate from the decay of a Higgs boson, whilst the latter refers to events which pass event selection, but in which not all leptons are prompt. These misidentified leptons can originate from

secondary decays of hadrons in jets or the hadrons themselves can be misidentified as leptons.

The contributions to the irreducible background are ZZ^* production, $t\bar{t}+Z$, ZZZ, WWZ and WZZ, which are all simulated using MC. These last three are collectively referred to as VVV. ZZ^* production is simulated by two MC samples separately for the production modes of quark fusion and gluon fusion which are abbreviated $qqZZ$ and $ggZZ$ respectively. The yield of each process is estimated using the MC.

The contributions to reducible background considered are $Z+$ jets, WZ and $t\bar{t}$, the yields for which are estimated using data-driven techniques in dedicated control regions [38].

5.3 Object requirements and event selection

Events are required to contain two same-flavour opposite-sign dilepton pairs (either electrons or muons). The requirements on these objects are summarised below.

5.3.1 Muons

Muons are required to fall in the acceptance range of the Muon System (MS), $|\eta|< 2.7$ [39]. The reconstruction of muons begins from reconstructing tracks from hits in the MS. These are extrapolated to the inner detector and any inner detector track falling within a cone of the extrapolated track is matched to the muon object. These tracks are fitted, accounting for energy loss in the calorimeters. Muons falling outside the full coverage of the inner detector ($2.5 \leq |\eta| < 2.7$) are reconstructed using only information from the MS, which can be matched to inner detector tracklets with only silicon hits, should these exist.

Loose identification requirements are used. Minimum hit requirements are set for muons with reconstructed inner detector tracks. For those outside the inner detector coverage, a minimum requirement of hits in the MDT and CSC components is used. Muon candidates with low quality tracks, missing measurements or poor agreement between MS and inner detector tracks are rejected. Muons are required to be isolated in both the calorimeter and inner detector. Muons are also required to have $p_{\mathrm{T}} > 5$ GeV and in the case of events with four muons, at least three must have associated inner detector tracks, i.e. $|\eta|< 2.5$. All muons are required to have have transverse impact parameter $|d_0| < 1$ mm, to reduce contamination from cosmic rays.

5.3.2 Electrons

The electron reconstruction, calibration and identification procedures are described in Chapter 4. However, note that in this analysis, an older method for clustering is used that relies on a sliding window approach rather than superclusters [32]. The Loose identification operating point is used, which rejects the most poorly reconstructed electrons. Electrons are also required to be isolated in both track and calorimeter and have $E_{\mathrm{T}} > 7$ GeV.

5.3.3 Event selection

The four leptons in an event, referred to as a quadruplet, are required to pass single-lepton, dilepton and trilepton triggers [6] and have a longitudinal impact parameter with respect to the beam spot $|z_0 \sin\theta| < 0.5$ mm [38]. The leptons are ranked by decreasing p_{T}, the highest, second highest and third highest p_{T} leptons are required to satisfy $p_{\mathrm{T}} > 20$, 15 and 10 GeV respectively.

The same-flavour, opposite-sign dilepton pair with invariant mass closest to the mass of the Z boson is referred to as the *leading pair*, with an invariant mass of m_{12}. The remaining pair of leptons have an invariant mass of m_{34} and are referred to as the *subleading pair*. Additional selections are placed upon the values of m_{12} and m_{34}, namely $50 < m_{12} < 106$ GeV and $m_{\mathrm{threshold}} < m_{34} < 115$ GeV. $m_{\mathrm{threshold}}$ depends on $m_{4\ell}$, which takes a value of 12 GeV for $m_{4\ell} < 140$ and 50 GeV for $m_{4\ell} > 190$ GeV, rising linearly in between. It is also required that all same flavour leptons in a quadruplet are separated by $\Delta R > 0.1$ and opposite flavour leptons by $\Delta R > 0.2$. A summary of the event selection is shown in Table 5.1.

Leptons and jets	
Muons:	$p_{\mathrm{T}} > 5$ GeV, $\lvert\eta\rvert < 2.7$
Electrons:	$p_{\mathrm{T}} > 7$ GeV, $\lvert\eta\rvert < 2.47$
Jets:	$p_{\mathrm{T}} > 30$ GeV, $\lvert y\rvert < 4.4$
Jet–lepton overlap removal:	$\Delta R(\mathrm{jet}, \ell) > 0.1\,(0.2)$ for muons (electrons)
Lepton selection and pairing	
Lepton kinematics:	$p_{\mathrm{T}} > 20, 15, 10$ GeV
Leading pair (m_{12}):	SFOS lepton pair with smallest $\lvert m_Z - m_{\ell\ell}\rvert$
Subleading pair (m_{34}):	remaining SFOS lepton pair with smallest $\lvert m_Z - m_{\ell\ell}\rvert$
Event selection (at most one quadruplet per channel)	
Mass requirements:	50 GeV$< m_{12} < 106$ GeV and 12 GeV $< m_{34} < 115$ GeV
Lepton separation:	$\Delta R(\ell_i, \ell_j) > 0.1\,(0.2)$ for same- (different-)flavour leptons
J/ψ veto:	$m(\ell_i, \ell_j) > 5$ GeV for all SFOS lepton pairs
Mass window:	115 GeV $< m_{4\ell} < 130$ GeV

Table 5.1 Summary of the event selection for the measurement of m_{H} at 36 fb^{-1}. [38]

Channel	Higgs ($m_{\mathrm{H}} = 125$ GeV)	ZZ^*	Other backgrounds	Total expected	Observed
4μ	20.1 ± 1.6	9.8 ± 0.8	1.3 ± 0.3	31.2 ± 1.8	33
$4e$	10.6 ± 1.0	4.4 ± 0.4	1.3 ± 0.2	16.3 ± 1.1	16
$2e2\mu$	14.2 ± 1.1	7.1 ± 0.5	1.0 ± 0.2	22.3 ± 1.2	32
$2\mu2e$	10.8 ± 1.0	4.6 ± 0.5	1.4 ± 0.3	16.8 ± 1.1	21
Inclusive	56 ± 4	25.9 ± 2.0	5.0 ± 0.7	87 ± 5	102

Table 5.2 Expected and observed yields for events within $115 < m_{4\ell} < 130$ GeV using ATLAS early Run 2 data. The expected yields for signal are calculated assuming $m_{\mathrm{H}} = 125$. [38]

The expected and observed event yields in the region $115 < m_{4\ell} < 130$ GeV are shown in Table 5.2. The 4μ and $4e$ yields are very close to those expected, whilst a small excess of $2e2\mu$ and $2\mu2e$ events is observed. The total expected and observed yields agree to within 1.3 standard deviations.

5.4 Improvements to $m_{4\ell}$ resolution

The resolution on $m_{4\ell}$ is improved using two approaches. Firstly, the recovery of final state radiation (FSR) photons. These are photons emitted in the decay of Z bosons. Including these leads to a more accurate reconstruction of the original Z boson and therefore improves the $m_{4\ell}$ resolution.

FSR photons are divided into two categories: *collinear* and *noncollinear*. Collinear photons are defined as having $\Delta R_{\mathrm{cluster},\mu} \leq 0.15$, where $\Delta R_{\mathrm{cluster},\mu}$ is the ΔR between the lepton and the photon cluster, and have $E_{\mathrm{T}} > 1.5$ GeV. They must also be associated with a muon only as, due to the tight ΔR requirement, such photons radiated in $Z \to ee$ decays would be included in the electron cluster. Noncollinear photons can be associated with either an electron or a muon, have $\Delta R_{\mathrm{cluster},\mu} > 0.15$, $E_{\mathrm{T}} > 10$ GeV, and also must be isolated and satisfy tight identification criteria (see Section 4.3 for an explanation of identification in the case of electrons).

A maximum of one FSR photon is used per event with a preference for collinear [40]. If more than one collinear or noncollinear photon is recovered, only that with the highest E_{T} is used.

The second contribution to the improvement in $m_{4\ell}$ is the use of a Z mass constraint. As the leading lepton pair originates predominately from resonant Z decays, a constrained kinematic fit is used to constrain m_{12} to the Z lineshape. This works by considering the probability that a generated Z boson has a true mass m_Z^{true}, given it

is reconstructed as m_{12}. This is proportional to the probability to produce such a Z boson at this mass multiplied by the resolution function, i.e.

$$p(m_Z^{\text{true}}|m_{12}, \sigma_{m_{12}}) \propto p(m_{12}|m_Z^{\text{true}}, \sigma_{m_{12}}) BW(m_Z^{\text{true}}|m_Z, \Gamma_Z), \tag{5.1}$$

where the first term on the RHS is probability to observe a value of m_{12} given that the Z boson has a true mass of m_Z^{true} with a measurement uncertainty of $\sigma_{m_{12}}$. The second term comes from the fact that the true lineshape of the Z boson decay is a relativistic Breit-Wigner function with m_Z as the Z pole mass and Γ_Z being the decay width [41, 42]. The constraint $m_{12} = m_Z^{\text{true}}$ is then applied and a maximum likelihood fit is used to maximise 5.1 by changing the lepton momenta.

In the following sections, $m_{4\ell}$ shall refer to the four-lepton invariant mass after the application of FSR recovery and the Z mass constraint.

5.5 Category definitions

The mass measurement is performed in the range $110 < m_{4\ell} < 135$ GeV. Events are placed exclusively into 16 categories based on the four channels (4μ, $4e$, $2\mu 2e$ and $2e2\mu$) and the event's score from a BDT. This BDT, which will be referred to as BDT_{ZZ}, is used to discriminate between signal and background based on the event's kinematics. It is trained using ggF MC (all mass points) for signal and ZZ^* MC for background. The inputs to this are the p_{T} and pseudorapidity of the four lepton system, $p_{\text{T}}^{4\ell}$ and $\eta_{\text{T}}^{4\ell}$ respectively, and a kinematic discriminant, $KD(ZZ^*)$, which uses the matrix elements from the SM production of ggF and ZZ^*. Although this discriminant requires m_{H} as an input, it does not bias the mass measurement. The output of BDT_{ZZ} is between -1 and 1, with -1 corresponding to most background-like and 1 corresponding to most signal-like. The 16 categories are defined by sub-dividing each of the four channels into four evenly-sized BDT_{ZZ} score bins. These bins are (-1,-0.5), (-0.5,0), (0,0.5) and (0.5,1). Separate signal and background models are created in each category with each of these models being fitted simultaneously to data to determine m_{H}.

5.6 Background model

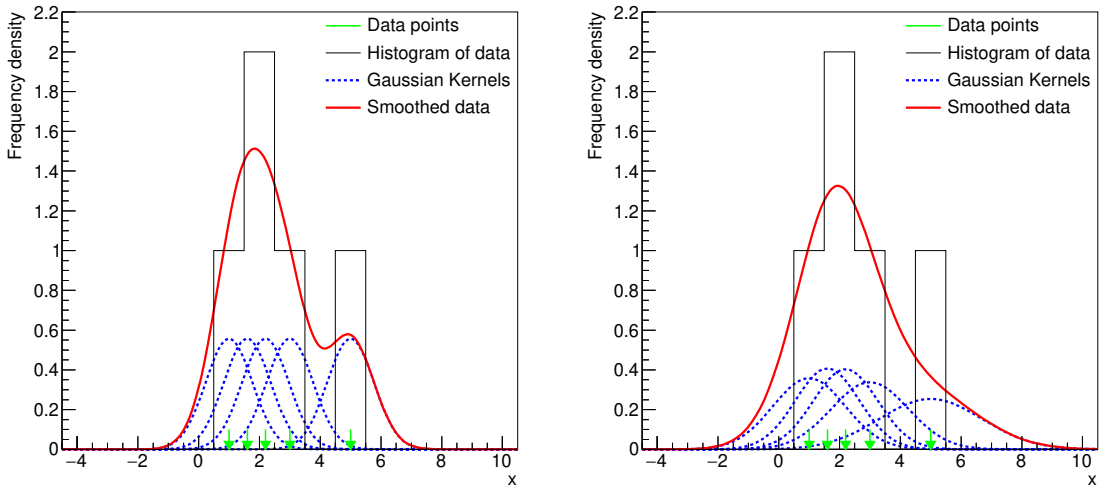
The shapes of the both the reducible and irreducible backgrounds are determined separately in each analysis category. To produce a PDF of the background, the $m_{4\ell}$

distributions from each MC sample for each process are smoothed using a technique of kernel density estimation (KDE smoothing) [36].

Kernel density estimation is a form of non-parametric description of a distribution. Most generally, a set of data points t_i is estimated by a PDF composed of the sum of *kernel* functions K as

$$\hat{f}_0(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - t_i}{h}\right), \tag{5.2}$$

where n is the number of data points and h is known as the *bandwidth* of the kernels. A common choice of kernel is a Gaussian, which is used in this analysis. A demonstration



(a) (b)

Fig. 5.1 Examples of KDE smoothing for the case of fixed-size kernels (a) and adaptive (b).

of this type of smoothing is shown in Figure 5.1a. To improve the estimate of the underlying distribution in regions where the data are fewer, it is desirable for the kernels to adapt. For this reason, the fixed-sized kernels of width h are replaced with variable width kernels with width

$$h_i = \rho \frac{h}{\sqrt{\hat{f}_0(t_i)}}, \tag{5.3}$$

where $\hat{f}_0(t_i)$ is that from Equation 5.2 and ρ is a constant referred to as the *smoothing parameter*. The adaptive description of the distribution is given by

$$\hat{f}_1(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_i}K\left(\frac{x-t_i}{h_i}\right). \tag{5.4}$$

An example of this type of smoothing is shown in Figure 5.1b.

Each of the components of the background is smoothed individually using this technique of adaptive KDE smoothing with Gaussian kernels.

5.7 Signal model

5.7.1 Per-event response method

The signal model used, known as the *per-event response* method, constructs the PDF of the signal lineshape from considering the individual energy responses of each lepton in each event. One key advantage of this approach over other methods, such as template fit methods (see Section 5.7.2), is that in per-event methods, events which are better measured have a greater impact on the fit. The lepton energy response is defined as

$$x_i = \frac{E_i^{\text{reco}} - E_i^{\text{truth}}}{E_i^{\text{truth}}}, \tag{5.5}$$

where E_i^{truth} and E_i^{reco} are respectively the true and reconstructed energies of lepton i in an event [41]. This is considered to be a random variable distributed as $P(x_i)$. By convolving the $P(x_i)$ for each of the four leptons in an event, the PDF $P(m_{4\ell}^{\text{reco}} - m_{4\ell}^{\text{truth}})$, where $m_{4\ell}^{\text{reco}}$ and $m_{4\ell}^{\text{truth}}$ are the reconstructed and true values of $m_{4\ell}$, can be determined. It is desirable for this PDF to be analytical for ease of calculation, so for this reason, $P(x_i)$ are modelled as the weighted sum of three Gaussians. The shape of these is determined by fitting to the simulated energy response distributions of electrons and muons separately in bins of leading/sub-leading Z, $|\eta|$ and p_{T}. $P(m_{4\ell}^{\text{reco}} - m_{4\ell}^{\text{truth}})$ therefore is the sum of $3^4 = 81$ Gaussians. Such a large number of Gaussians per event would be very computationally intensive, so the number is reduced using a system of Gaussian mixture reduction. To do this, a "distance" between any two Gaussians in the mixture is defined. The distance between any two Gaussians i and j is given by

$$D_{ij} = (\mu_i - \mu_j)^2\left(\frac{1}{\sigma_i^2} + \frac{1}{\sigma_j^2}\right) + \frac{\sigma_i^2}{\sigma_j^2} + \frac{\sigma_j^2}{\sigma_i^2}, \tag{5.6}$$

where μ_i and μ_j denote the means of Gaussians i and j respectively and σ_i and σ_j are their standard deviations. This quantity is known as the *Kullback-Leibler distance* for two Gaussians [41]. The pair of Gaussians with the smallest D_{ij} is found and replaced with a single Gaussian. The weight of this new Gaussian is given by

$$w' = w_i + w_j, \tag{5.7}$$

where w_i and w_j are the weights of the original Gaussian. The mean and variance are given by

$$\mu' = \frac{w_i\mu_i + w_j\mu_j}{w_i + w_j} \tag{5.8}$$

and

$$\sigma'^2 = \frac{w_i\sigma_i^2 + w_j\sigma_j^2}{w_i + w_j} + \frac{w_i w_j}{(w_i + w_j)^2}(\mu_i - \mu_j)^2 \tag{5.9}$$

respectively. This process preserves the mean and variance of the original 81 Gaussian distribution, but not higher moments. Figure 5.2 shows a visualisation of the procedure

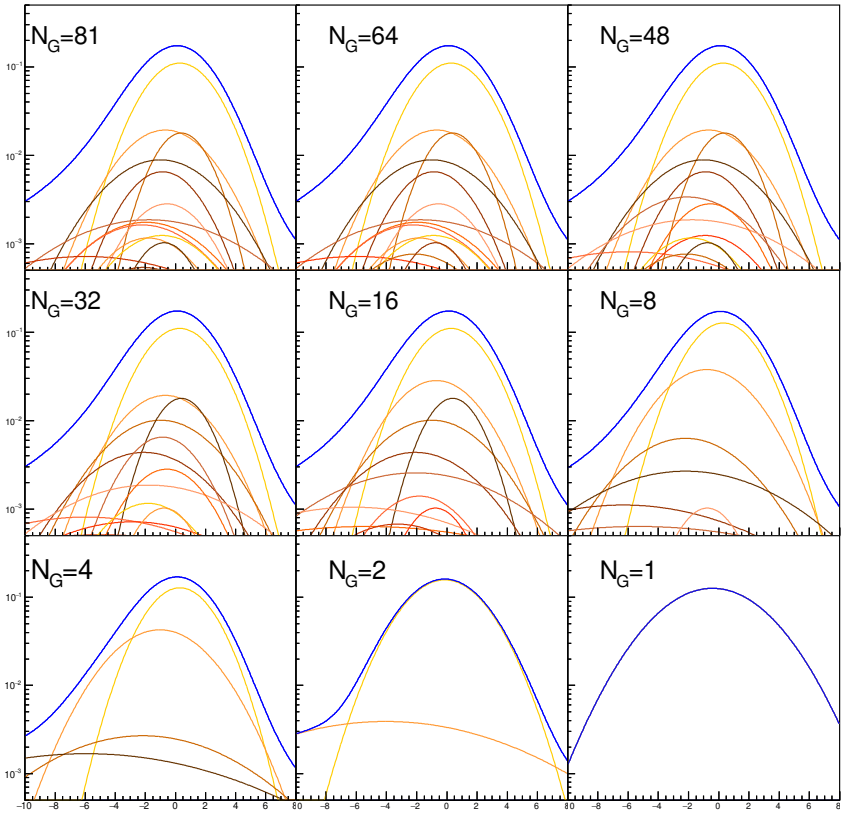


Fig. 5.2 A visualisation of the Gaussian mixture reduction procedure. [41]

on an MC event. By this procedure, the number of Gaussians per-event is reduced to 4. The reason for this number is that four Gaussians is the minimum number required to sufficiently describe an event. Using 3 Gaussians, clear discrepancies are seen and using 5 sees little improvement.

The distribution $P(m_{4\ell}^{\text{reco}} - m_{4\ell}^{\text{truth}})$ is related to the observed $m_{4\ell}$ distribution, $P(m_{4\ell}^{\text{reco}})$, by the convolution

$$P(m_{4\ell}^{\text{reco}}) = P(m_{4\ell}^{\text{reco}} - m_{4\ell}^{\text{truth}}) * P(m_{4\ell}^{\text{truth}}), \tag{5.10}$$

where $P(m_{4\ell}^{\text{truth}})$ is the distribution of the "true" values of $m_{4\ell}$ (i.e. without detector smearing). This is modelled as a Breit-Wigner distribution as

$$P(m_{4\ell}^{\text{truth}}) = BW(m_{4\ell}^{\text{truth}}; m_H, \Gamma_H), \tag{5.11}$$

where Γ_H is the natural width of the Higgs boson. As Γ_H 1 MeV and the width of the reconstructed Higgs peak is 1 GeV, this is in practice a convolution with a delta function [43]. The model is not sensitive to the value of Γ_H and so it is not a free parameter in the fit. Using this, Equation 5.10 can be fitted to the distribution of $P(m_{4\ell}^{\text{reco}})$ from data to determine m_{H}.

5.7.2 Template method

A method of creating unbinned templates of the $m_{4\ell}$ distribution is used as a cross check to the above. This method was used as the primary method in the ATLAS m_{H} measurement at 25 fb^{-1} [37]. In contrast to the per-event method, here the distribution of $P(m_{4\ell}^{\text{reco}})$ vs m_{H} is determined directly (rather than by using lepton response distributions), which is done by the KDE smoothing of the $m_{4\ell}$ distribution of MC signal samples at various different simulated Higgs boson masses. In order to obtain a continuous model of $m_{4\ell}$ vs m_{H}, the templates at each mass are interpolated using a B-spline interpolation. Figure 5.3 shows this interpolation in the 4μ channel.

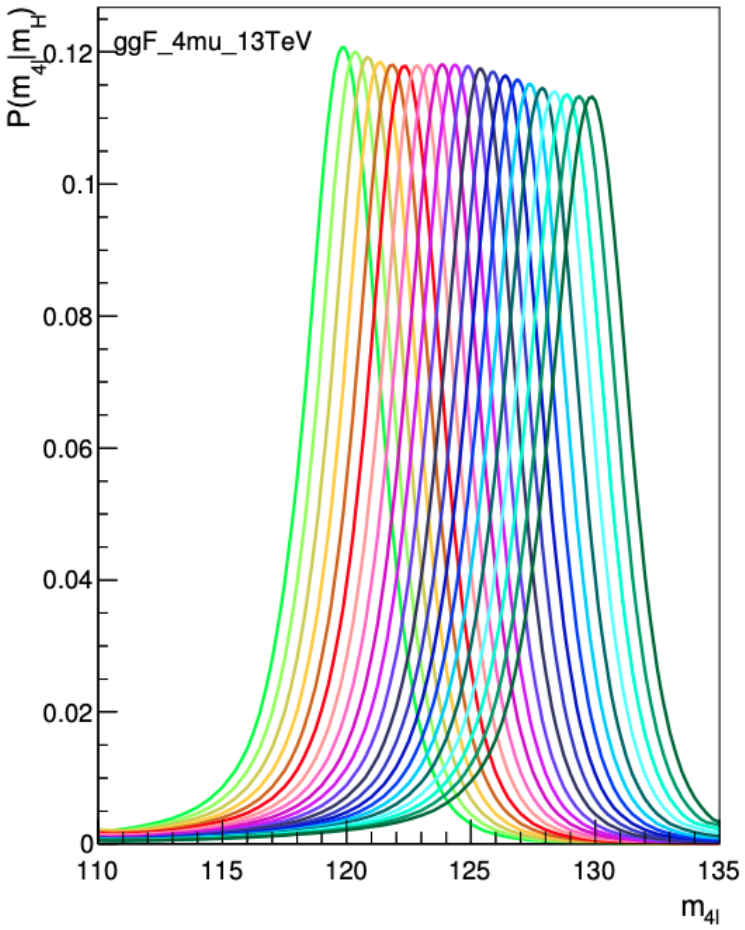


Fig. 5.3 Example of the morphing of template $m_{4\ell}$ distributions in the 4μ channel. [41]

5.8 Results

The value of m_{H} is calculated by performing a simultaneous fit to the data across all 16 analysis categories. From this, the mass of the Higgs boson is determined to be

$$m_{\mathrm{H}} = 124.79 \pm 0.36 \text{ (stat.)} \pm 0.05 \text{ (syst.) GeV},$$

The dominant sources of systematic uncertainty here arise from the electron energy scale uncertainty and the muon momentum scale uncertainty. The projection of the fit across all 16 categories as well as the likelihood profiles for the four decay channels are shown in Figure 5.4.

This in agreement the measurement of m_{H} in the $H \to ZZ^* \to 4\ell$ from CMS over the same data taking period of 125.26 ± 0.25 (stat.) ± 0.08 (syst.) GeV.

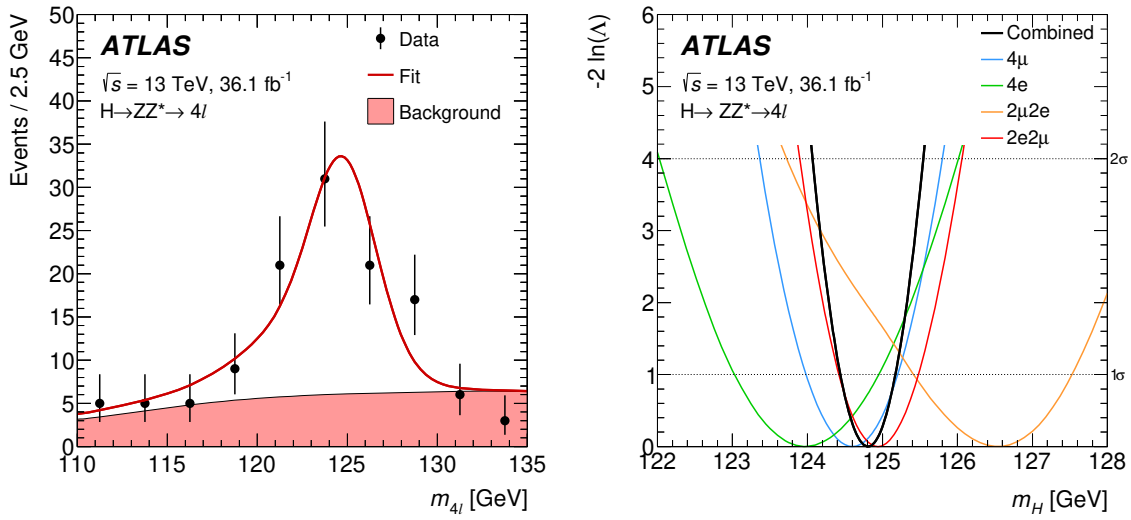


Fig. 5.4 Results for the Higgs boson mass measurement at 36 fb^{-1} in the $H \to 4\ell$ channel. Left, the projection of the fit from all 16 categories and right, the profile likelihood of the fit to each channel. [6]

The other channel using which a mass measurement can be performed is the $H \to \gamma\gamma$ channel. Whilst having more signal events than $H \to ZZ^* \to 4\ell$, two of the drawbacks are that the systematic uncertainties are larger and the proportion of background is higher. The ATLAS measurement in this channel using data collected during LHC early Run 2 is

$$m_{\mathrm{H}} = 124.93 \pm 0.21 \text{ (stat.)} \pm 0.34 \text{ (syst.) GeV},$$

which is in good agreement with the $H \to ZZ^* \to 4\ell$ measurement [6].

The ATLAS $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ results from early Run 2 are combined to give

$$m_{\mathrm{H}} = 124.86 \pm 0.18 \text{ (stat.)} \pm 0.20 \text{ (syst.) GeV},$$

in agreement with the ATLAS combined LHC Run 1 result of 125.38 ± 0.37 (stat.) $\pm$ 0.18 (syst.) GeV [6].

The ATLAS combined Run 1 and Run 2 results are combined into a single measurement. The result of this combination is

$$m_{\mathrm{H}} = 124.97 \pm 0.16 \text{ (stat.)} \pm 0.18 \text{ (syst.) GeV},$$

which is in good agreement with the combination of ATLAS and CMS results from Run 1 of 125.09 ± 0.16 (stat.) $\pm$ 0.18 (syst.) GeV [37, 44, 45]. The comparison of these results is shown in Figure 5.5.

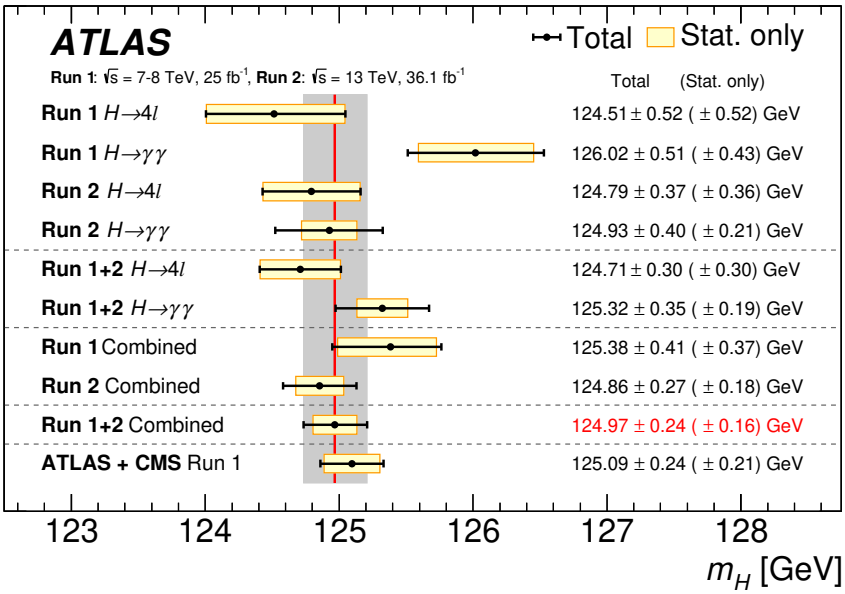


Fig. 5.5 Combination of ATLAS Run 1 and early Run 2 measurements of m_{H} in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels. This is compared to the combined ATLAS + CMS result from Run 1, with which good agreement is seen. [6]

Chapter 6

Standard model Higgs Boson mass measurement in $H \to 4\ell$ at 139 fb^{-1}

This chapter details the standard model Higgs boson mass measurement in the $H \to ZZ^* \to 4\ell$ channel using the ATLAS full LHC Run 2 dataset, corresponding to an integrated luminosity of 139 fb^{-1} from pp collisions at a centre of mass energy of 13 TeV. It contains significant contributions by the author.

6.1 Analysis strategy

For this measurement, the per-event response method described in the previous chapter is replaced by an analytic model incorporating per-event errors [46]. Whilst the previously used per-event response model provides an excellent description of the detector response per event, it is computationally intensive, increasingly so with increasing numbers of events. Unlike the per-event response method, which begins with a description of each lepton in an event from which a PDF of $m_{4\ell}$ is built, an analytic model attempts to construct a PDF to describe the average lepton, without considering the underlying lepton kinematics. A Double-sided Crystal Ball (DCB) PDF (see Section 6.2.1) is used to model the $m_{4\ell}$ spectrum. The σ parameter, which corresponds to the width of the DCB, represents the average resolution of an event. To account for varying event resolution, a DCB is fitted conditional to σ, with the per-event value of σ being calculated using machine learning techniques (Section 6.2.2). The benefits of an analytic model are that it is computationally less intensive than a per-event response method whilst extensions to incorporate additional observables (e.g. BDT discriminant and per-event error) are much easier than in the case of a template method.

The event selection remains largely unchanged from Section 5.3, but the range of the fit has been expanded to $105 < m_{4\ell} < 160$ GeV. This allows the normalisation of the background using side-bands (Section 6.4), which are defined as $105 < m_{4\ell} < 115$ and $130 < m_{4\ell} < 160$. Other changes are that all leptons are required to be separated by $\Delta R > 0.1$, whereas previously, the requirement was stricter for opposite-sign leptons. The muon isolation is also updated to use variables which perform better at higher pile-up. The Z mass constraint and FSR recovery procedures remain unchanged.

6.2 Signal model

All simulated MC samples are updated to take into account the increase in pile-up at the LHC. As previously, the ggF and VBF signal production modes are simulated at several different values of m_{H} for use in building the signal model. Other signal processes are used for validation of the signal model and determining expected event yields. New for this analysis, Higgs boson production in association with: a single top quark (tH), a single top quark and a W boson (tWH) and a pair of bottom quarks ($b\bar{b}H$) are also simulated. A list of all signal samples used is shown in Table 6.1.

6.2.1 Analytic parameterisation

The analytic signal model is fitted to the MC using an unbinned-maximum likelihood fit [47]. This method is used as information is not lost in binning of the data and parameters of interest tend to be Gaussian distributed in the asymptotic limit [48]. Suppose a PDF, P, with a vector of parameters $\mathbf{a}$ is chosen to fit data points x_i. The expression for the likelihood is given by

$$L(x_i) = \prod_i^n P(x_i; \mathbf{a}), \tag{6.1}$$

where n is the number of data points, i.e. the number of observed events. The optimal values of the parameters, $\mathbf{a}$, are those which maximise L. The likelihood is implemented using RooFit [49]. The normalisation is incorporated into the likelihood fit with a Poisson term for the number of events observed

$$L(x_i) = e^{-N} \frac{N^n}{n!} \prod_i^n P(x_i; \mathbf{a}), \tag{6.2}$$

Production mode	m_{H} [GeV]	N_{MC}	N_{expt}
ggF	123.0	174542	151.1
	124.0	1079046	166.8
	124.5	1093500	175.7
	125.0	1534376	186.1
	125.5	1127239	195.7
	126.0	1103111	206.0
	127.0	196927	227.3
VBF	123.0	364626	13.6
	124.0	648136	15.2
	125.0	1925600	17.0
	125.5	644594	17.9
	126.0	3773226	18.8
	127.0	469416	20.9
W^+H	124.0	44661	2.64
	124.5	45911	2.80
	125.0	73787	2.94
	125.5	46360	3.02
	126.0	47424	3.20
W^-H	124.0	50313	1.90
	124.5	50612	1.96
	125.0	81700	2.05
	125.5	52051	2.18
	126.0	52700	2.25
ZH	124.0	52753	3.05
	124.5	53484	3.21
	125.0	162699	3.36
	125.5	54346	3.51
	126.0	54883	3.66
$ggZH$	125.0	137630	0.61
$t\bar{t}H,\ t\bar{t} \rightarrow$ all hadronic	125.0	152823	0.75
$t\bar{t}H,\ t\bar{t} \rightarrow$ semi-leptonic	125.0	187536	0.78
$t\bar{t}H,\ t\bar{t} \rightarrow$ dilepton	125.0	127423	0.22
tH	125.0	50932	0.29
tWH	125.0	25011	0.06
$b\bar{b}H$	125.0	34859	2.06
$b\bar{b}H$	125.0	213601	-0.16

Table 6.1 A list of all the simulated signal processes. MC simulations for the two leading processes, ggF and VBF are simulated at several different mass points as these are used for building the model. The sub-leading processes are only used in validation of the model and determining the expected event yield. N_{MC} is the number of unweighted MC events which pass selection and N_{expt} is the expected number of events at 139 fb^{-1} in the $105 < m_{4\ell} < 160$ GeV range.

where N is the expected number of events. For computational reasons, this is usually expressed in the equivalent logarithmic form

$$\ln L(x_i) = \sum_i^n \ln P(x_i; \mathbf{a}) + n \ln N - N + \text{constant}, \tag{6.3}$$

which is referred to as the *log likelihood*.

Choice of PDF

The PDF used is a Double-sided Crystal Ball function (DCB) [50, 51, 52]. This is a function with a Gaussian core and two independent power-law tails. The function defined as

$$\begin{aligned} P(m_{4\ell}) &= DCB\left(m_{4\ell}; \mu, \sigma, \alpha_{\text{low}}, n_{\text{low}}, \alpha_{\text{high}}, n_{\text{high}}\right) \\ &= \begin{cases} \left(\frac{n_{\text{low}}}{\alpha_{\text{low}}}\right)^{n_{\text{low}}} \cdot e^{-\frac{\alpha_{\text{low}}^2}{2}} \cdot \left(\frac{n_{\text{low}}}{\alpha_{\text{low}}} - \alpha_{\text{low}} - \frac{m_{4\ell}-\mu}{\sigma}\right)^{-n_{\text{low}}}, & \left(\frac{m_{4\ell}-\mu}{\sigma}\right) < -\alpha_{\text{low}} \\ G\left(m_{4\ell}; \mu, \sigma\right), & -\alpha_{\text{low}} < \left(\frac{m_{4\ell}-\mu}{\sigma}\right) < \alpha_{\text{high}} \\ \left(\frac{n_{\text{high}}}{\alpha_2}\right)^{n_{\text{high}}} \cdot e^{-\frac{\alpha_{\text{high}}^2}{2}} \cdot \left(\frac{n_{\text{high}}}{\alpha_{\text{high}}} - \alpha_{\text{high}} + \frac{m_{4\ell}-\mu}{\sigma}\right)^{-n_{\text{high}}}, & \text{else} \end{cases} \end{aligned}, \tag{6.4}$$

where G is a Gaussian with a peak position μ and standard deviation σ. n_{low} and n_{high} are the exponents of the low and high tails respectively and α_{low} and α_{high} describe where the low and high power law tails begin respectively.

A second function is also investigated, the sum of a Crystal Ball function (CB) with a single power law lower tail and a Gaussian, denoted collectively as CBG. The function is

$$P(m_{4\ell}) = f_{CB} \cdot \text{CB}(m_{4\ell}; \mu, \sigma_{CB}, \alpha, n) + (1 - f_{CB}) \cdot \text{G}(m_{4\ell}; \mu + \Delta\mu, \sigma_G), \tag{6.5}$$

where μ is position of the peak of the CB, $\Delta\mu$ is the difference between the positions of the peaks of the Gaussian and the CB, f_{CB} is the relative fraction of the CB, σ_{CB} and σ_G are the widths of the CB and Gaussian respectively, n is the exponent of the power law tail of the Crystal Ball and α describes where this tail begins. Both functions are fitted separately to each of the four channels, 4μ, $4e$, $2\mu 2e$ and $2e2\mu$. To determine the quality of the fits, χ^2 per degree of freedom is used. To do this, the MC and the fitted

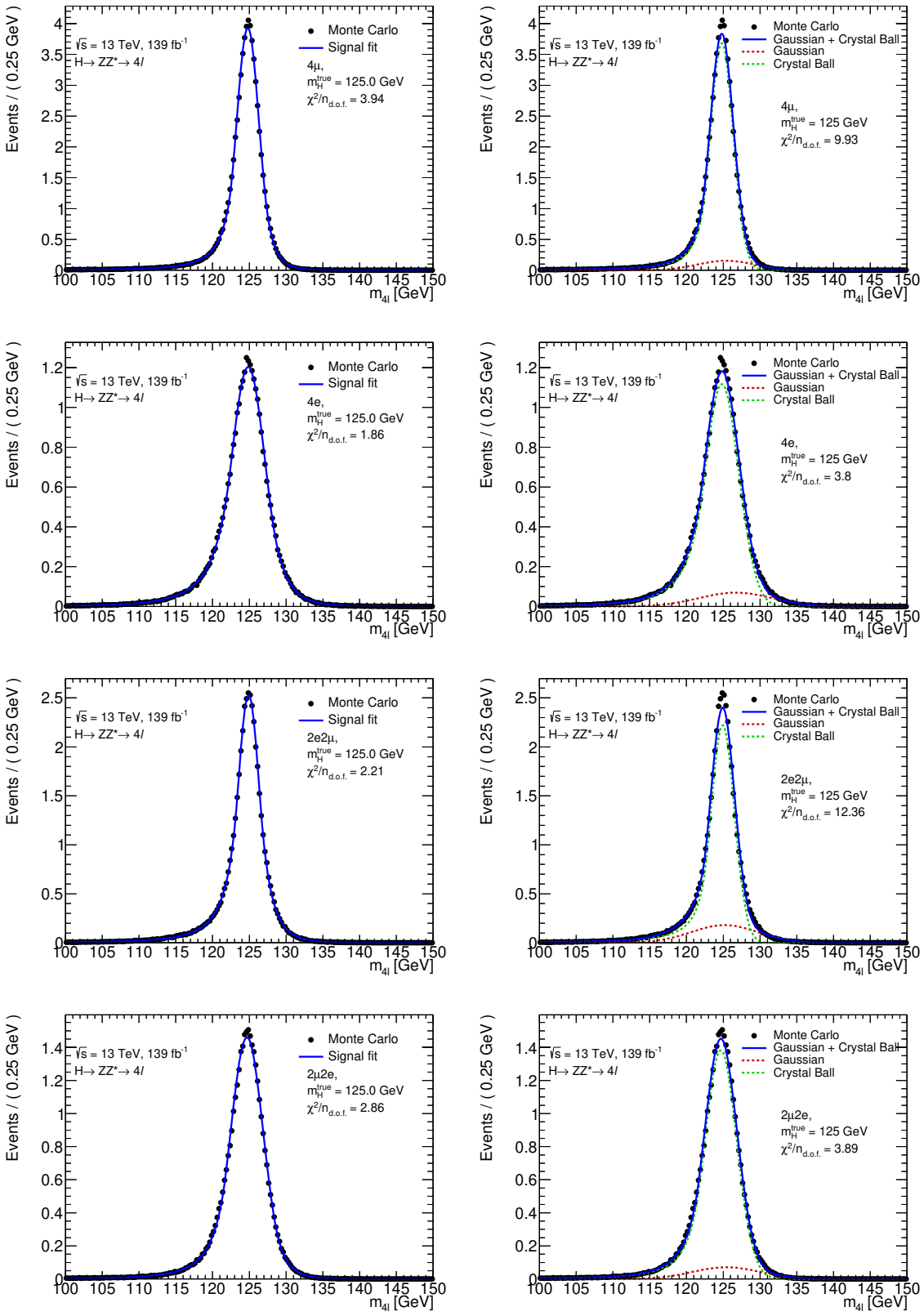


Fig. 6.1 Fits of DCB (left column) and CBG (right column) models to Monte Carlo simulations for $m_{\mathrm{H}} = 125$ GeV, ggF production mode in each channel shown in a linear scale. The DCB model fits better in all channels.

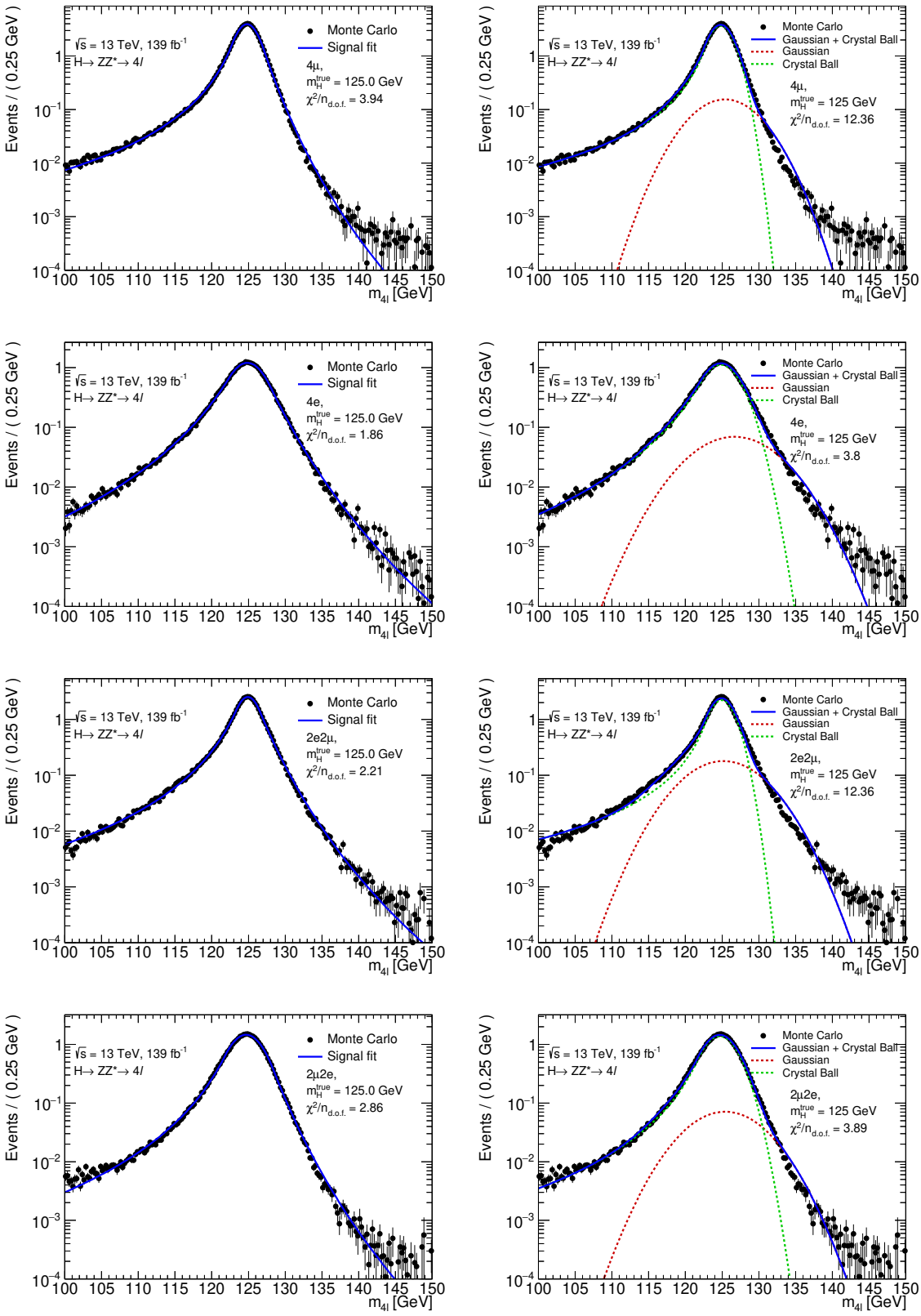


Fig. 6.2 Fits of DCB (left column) and CBG (right column) models to Monte Carlo simulations for $m_{\mathrm{H}} = 125$ GeV, ggF production mode in each channel shown in a log scale. The DCB model fits better in all channels.

function must be plotted in a binned manner and the value of χ^2 is given by

$$\chi^2 = \sum_i^N \frac{(O_i - E_i)^2}{\sigma_i^2}, \tag{6.6}$$

where O_i and E_i are the observed and expected (fitted) number of events in a bin i, σ_i is the uncertainty on O_i and N is the number of bins. The number of bins is chosen to be 200 and σ_i^2 is calculated using a sum of weights squared method. Figures 6.1 and 6.2 show a comparison of the CBG and DCB models when fitted to $H \to 4\ell$ MC for $m_{\mathrm{H}} = 125$ GeV, ggF production mode in each channel, using unbinned maximum likelihood fitting with the $\chi^2/n_{\mathrm{d.o.f.}}$ shown. The DCB model is a better fit for the MC than the CBG model in each channel. Although the values of $\chi^2/n_{\mathrm{d.o.f.}}$ are still poor, this is mainly due to the tails not fitting as well as the core. This however has little effect on the mass measurement as the parameter which affects the measurement of m_{H} the most is μ, which is not affected by the tails. As the DCB fits better in each channel, it is the model used in this analysis.

Choice of fit range

One of the challenges of fitting Double-sided Crystal Ball functions is the instability of the tail parameters n_{low}, n_{high}, α_{low} and α_{high}. That is, a small fluctuation in the dataset can cause a large change in one of these four parameters. Whilst this instability is impossible to avoid completely, its effect can be reduced by the choice of fit range[1].

The low tail parameters and high tail parameters are investigated separately. A fit is performed to each of the four channels in the $m_{4\ell}$ range $110 < m_{4\ell} < 135$ GeV to the ggF $m_{\mathrm{H}} = 125$ GeV mass point. This is taken to be the nominal case. First, the high tail parameters α_{high} and n_{high} are investigated. To do this, another fit is performed with the fitting window changed to $110 < m_{4\ell} < 136$ GeV. The parameters of the DCB are initially set to their nominal values and allowed to float. This process is repeated, each time increasing the size of the window by 1 GeV up to 150 GeV. Values of the parameters α_{high} and n_{high} are shown vs the upper limit of the window, x_{max} in Figure 6.3. The n_{high} and α_{high} parameters stabilise as the window is enlarged.

The same test is performed for the n_{low} and α_{low}. For this, the window is started at the nominal size and the lower bound is reduced in 1 GeV steps until 100 GeV is reached. The result of this is shown in Figure 6.4. As with the upper limit of the

[1]Whilst data events in the range $105 < m_{4\ell} < 160$ GeV are used to determine m_{H}, this need not necessarily correspond to the range used to calibrate the signal model.

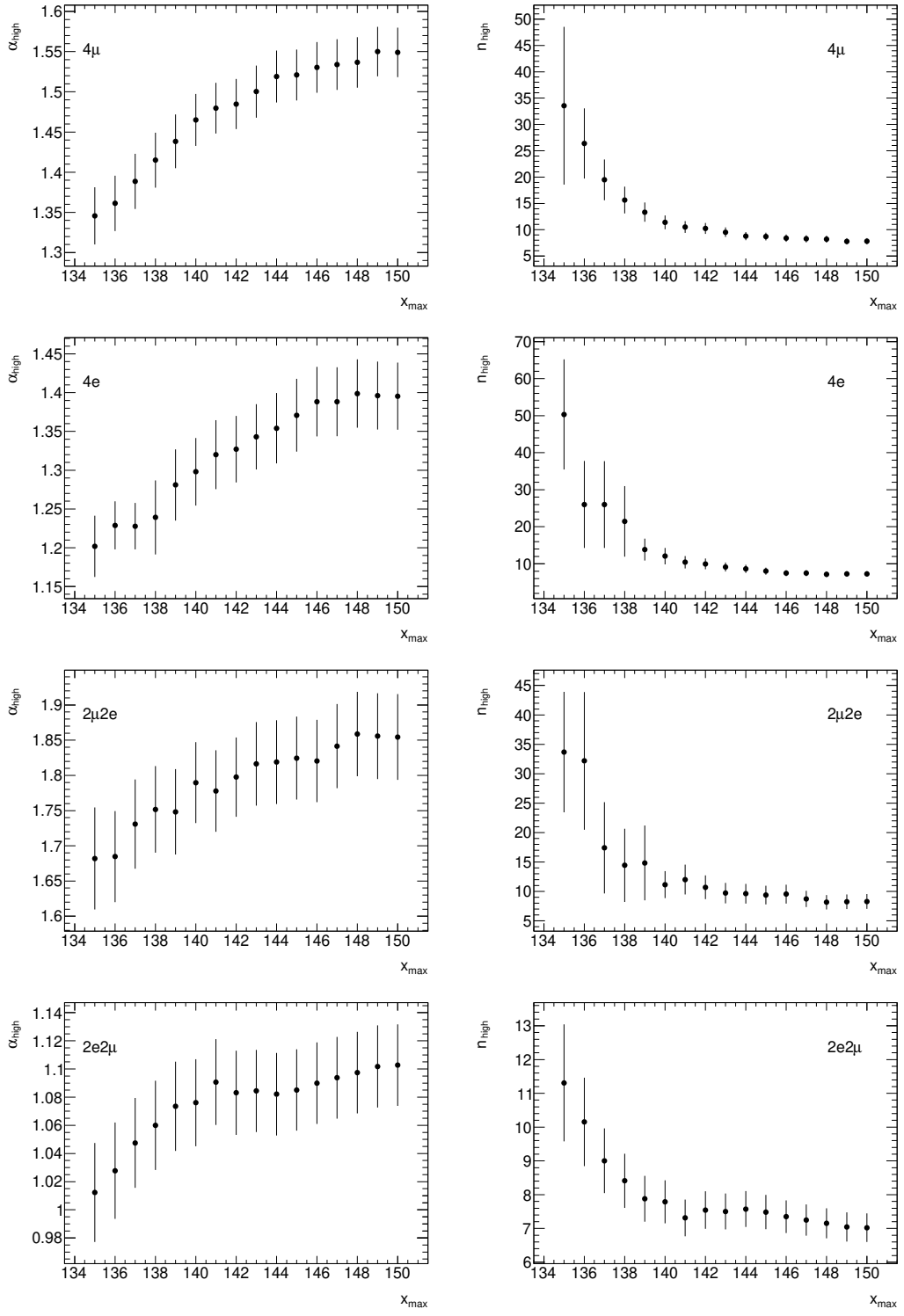


Fig. 6.3 Left: α_{high} vs x_{max} and right: n_{high} vs x_{max} for each channel. The α_{high} and n_{high} parameters tend to reach more stable values with increasing x_{max}.

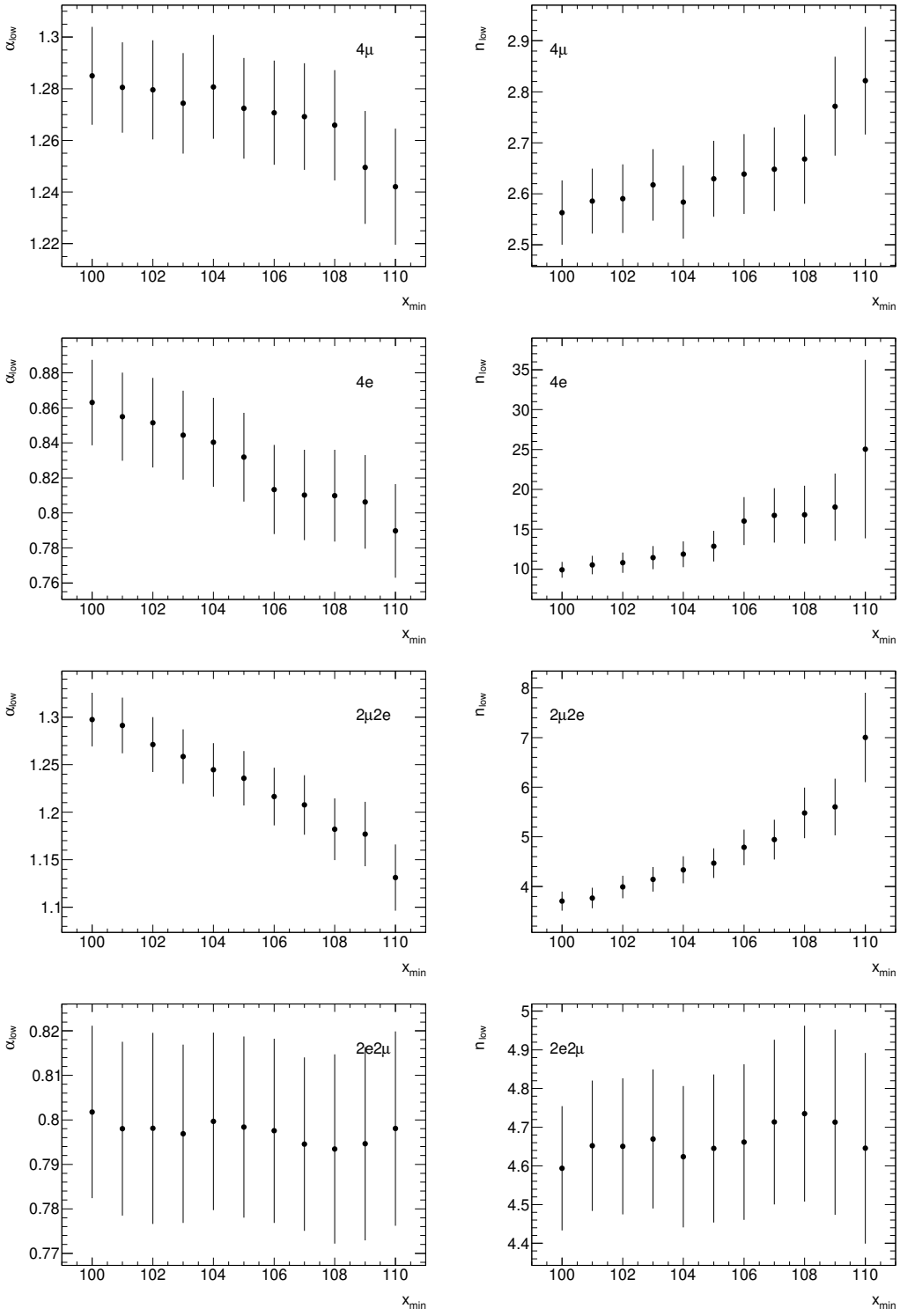


Fig. 6.4 Left: α_{low} vs x_{low} and right: n_{low} vs x_{min} for each channel. The α_{low} and n_{low} parameters tends to reach more stable values with decreasing x_{min}.

fit window, the fit stabilises as the window is enlarged. For these reasons, a window $100 < m_{4\ell} < 150$ GeV is used to calibrate the signal model.

Construction of the model

The general strategy is to build models at discrete values of m_{H}, using the MC samples simulating $H \to 4\ell$ events at several different mass points. This is used to build a continuous analytic description of the $m_{4\ell}$ spectrum vs m_{H}, which is fitted to the data. However, first a simple test is performed to verify the functionality of the DCB model as a method to determine m_{H}.

This simple test is done by randomly sampling the MC dataset for the $m_{\mathrm{H}} = 125$ GeV ggF mass point in the 4μ channel. From doing this, two equal sized MC datasets are created, which are referred to as the *calibration* and *validation* sets. Next a DCB is fitted to the calibration set with all parameters floating. For a test of the closure, the peak of the DCB, μ, is parametrised in terms of m_{H} as a simple offset, namely

$$m_{\mathrm{H}}^{\mathrm{measured}} = \mu + \Delta\mu, \tag{6.7}$$

where $m_{\mathrm{H}}^{\mathrm{measured}}$ is the measured value of m_{H}, $\Delta\mu$ is the difference between the value of μ and the true mass, $m_{\mathrm{H}}^{\mathrm{true}}$ in the calibration set. The model is validated by fitting the same DCB as in the previous step to the validation set. In this fit, only the μ parameter is allowed to float. $\Delta\mu$, and all other parameters are fixed to their values from the calibration set. Equation 6.7 is then used to calculate a value of $m_{\mathrm{H}}^{\mathrm{measured}}$. These fits are shown in Figure 6.5.

This procedure is repeated for each of the other three channels. A combined result is also produced by fitting simultaneously across all four channels. A summary of the results from this is shown in Figure 6.6.

The combined fit shows closure, so the simple model is valid for the case where $m_{\mathrm{H}} = 125$ GeV. However, parametrising as in Equation 6.7 assumes that m_{H} and μ follow a linear relationship with a slope of exactly 1. Figure 6.7 shows the variation in fitted μ vs m_{H} in each channel. From this it is clear that μ has a very strong linear relationship with m_{H}, but the slope is < 1, so a simple offset is inexact. Clearly, a linear parameterisation of m_{H} with μ is needed. The following linear parameterisation is used:

$$\mu = a_\mu(m_H - 125) + b_\mu + 125, \tag{6.8}$$

where a_μ is the gradient of the slope and b_μ is the offset. The 125 terms are to reduce the impact of the error in a_μ on μ. The approach used to implement this linear

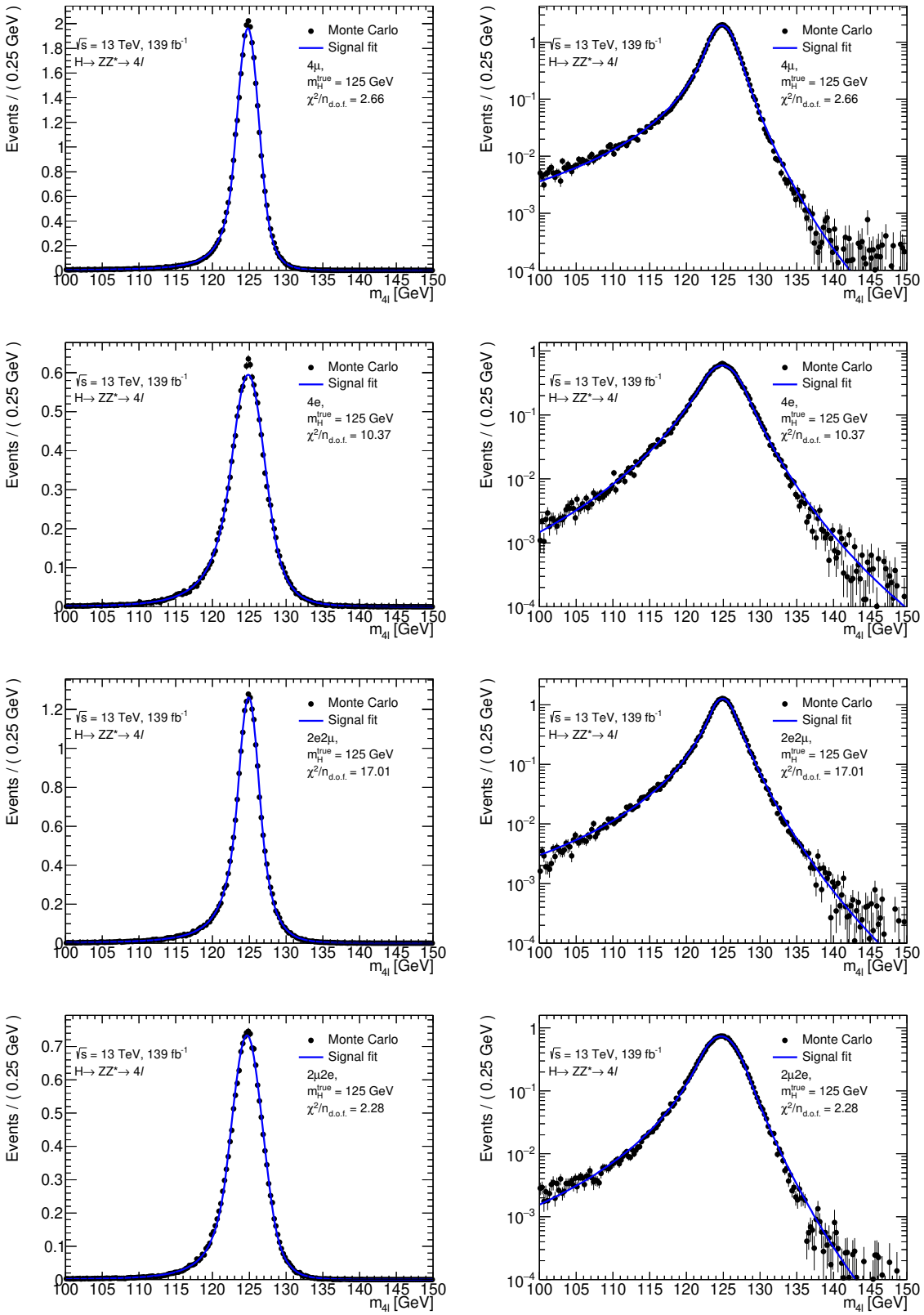


Fig. 6.5 Validation of the DCB parameterisation on each channel, shown in a linear and a log scale. The DCB is fitted to the second set with μ floating and all other parameters fixed. From these fits, Equation 6.7 is used to estimate m_{H}.

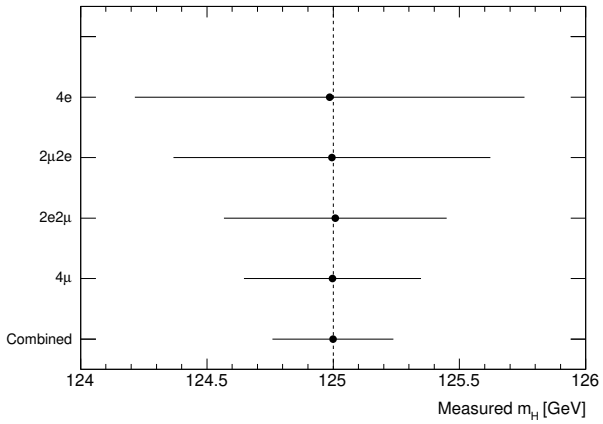


Fig. 6.6 The measured values of m_{H} from validation fits in each of the four channels as well as the combined fit across all channels. The uncertainties here reflect the expected statistics at 139 fb^{-1}.

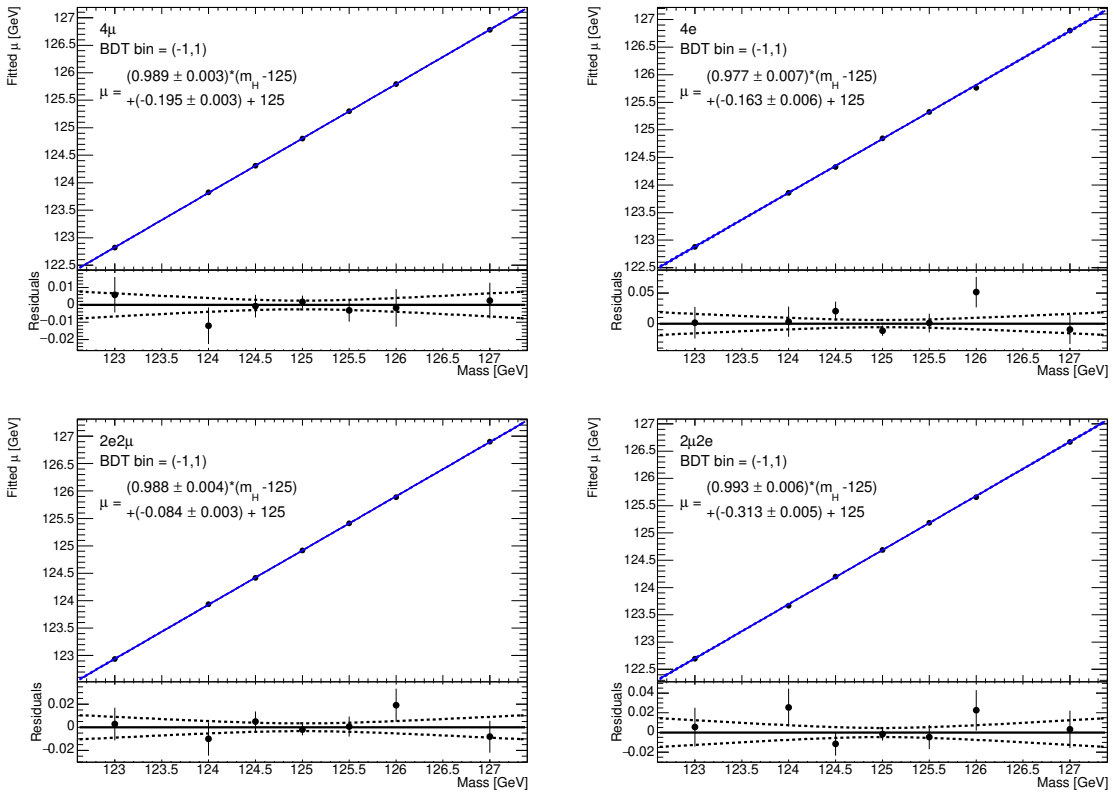


Fig. 6.7 Fitted μ vs true m_{H} from fitting a DCB model to MC mass points in the range 123 to 127 GeV for each channel. The μ parameter shows a strong linear relationship with m_{H}; however, the slope is not exactly 1 in each case.

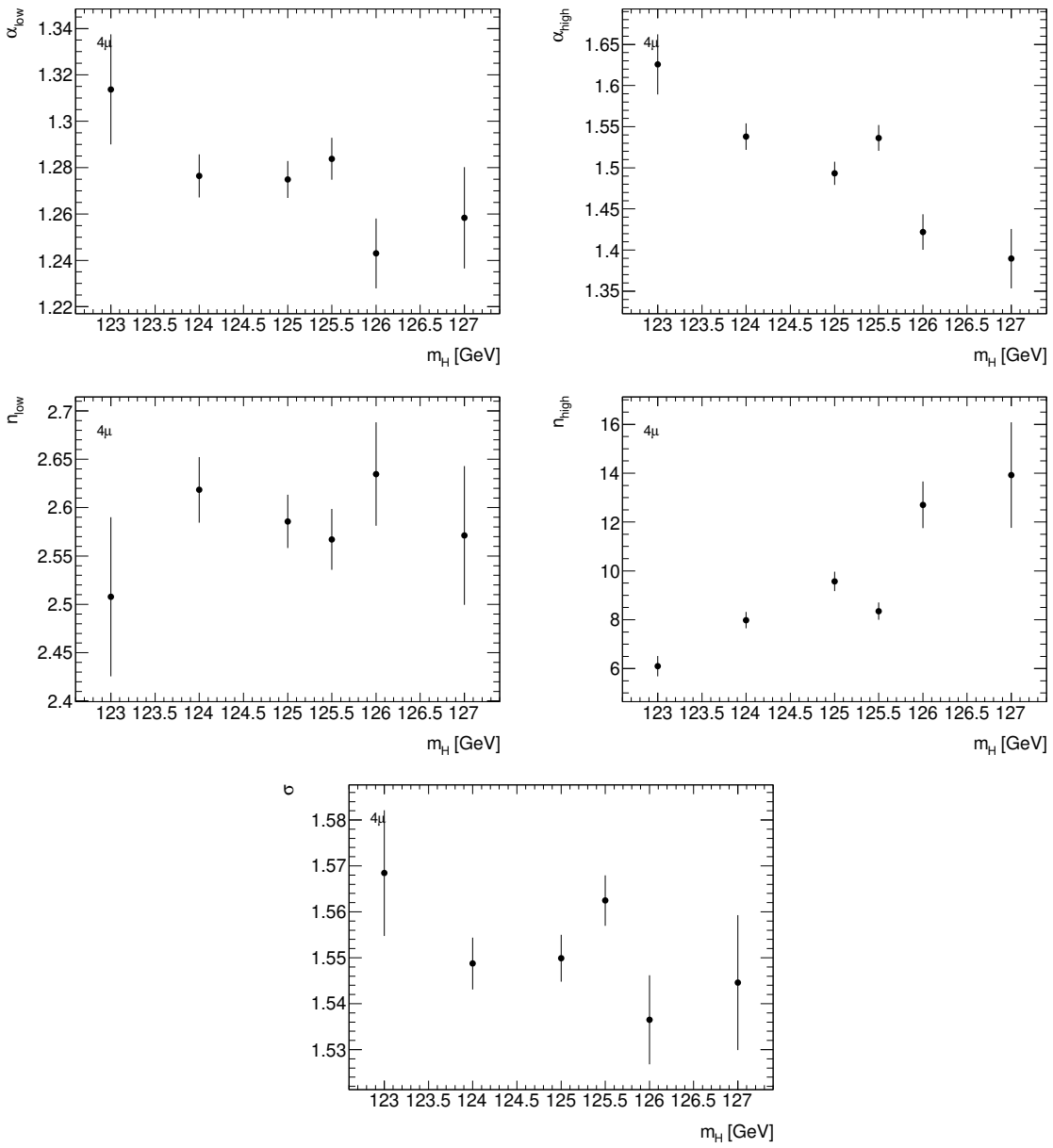


Fig. 6.8 All DCB parameters (except μ) vs m_{H} in the 4μ channel. The α_{high} and n_{high} parameters show some dependence on m_{H} whereas all other parameters show no strong dependence.

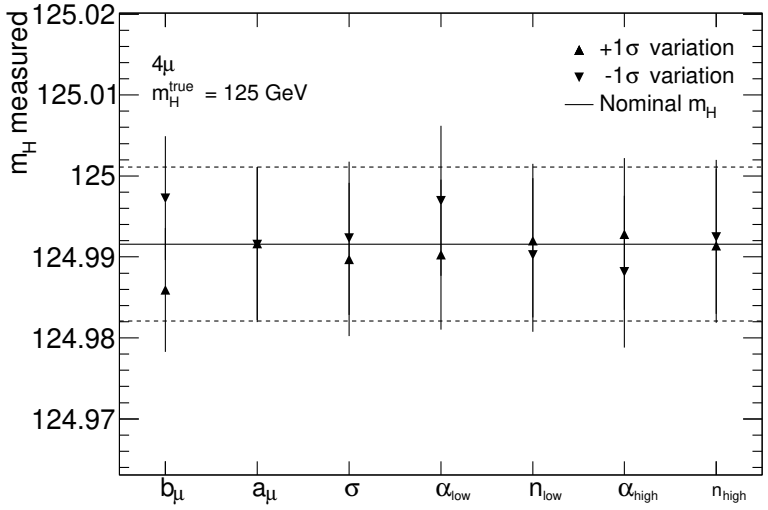


Fig. 6.9 Bias in m_{H} measurement in 4μ channel from varying each parameter of the model by $\pm 1\sigma$, where σ is its uncertainty. Starting from the nominal model, one parameter is increased by $+1\sigma$ from its nominal value, the model is fitted to the $m_{\mathrm{H}}^{\mathrm{true}}$ = 125 GeV MC set and the value of $m_{\mathrm{H}}^{\mathrm{measured}}$ is plotted. This process is repeated for the $+1\sigma$ case and for each parameter in the model. The value of $m_{\mathrm{H}}^{\mathrm{measured}}$ from the nominal model is shown as well as its uncertainty.

parameterisation into the model is to simultaneously fit one DCB to each mass point, each having common values of a_μ and b_μ. As Figure 6.8 shows, the n_{low}, α_{low} and σ parameters of the DCB in the 4μ channel show little dependence on m_{H}. It is also of note that the points on these graphs fluctuate a large amount, with the $m_{\mathrm{H}} = 125.5$ GeV being a particularly large outlier. For this reason, they are taken to be constant vs m_{H}. However, the parameters of the high tail show some dependence. Whilst it is possible to parametrise these parameters linearly, it is undesirable to do so, due to their instability. To investigate whether they can be kept constant with m_{H}, the effect of a systematic variation in these parameters is studied. Using the nominal model, one of the parameters is increased by the value of its uncertainty ($+1\sigma$), the model is fitted to the ggF MC simulation for $m_{\mathrm{H}} = 125$ GeV and a value of m_{H} is recovered. This process is repeated for the case where the parameter is changed by -1σ from its nominal value and for each other parameter of the model. The results from this test in the 4μ channel are shown in Figure 6.9. From this, the n_{high} and α_{high} parameters have very little effect on the value of m_{H} and so can be parametrised as constant vs m_{H}.

As in the previous analysis, to improve signal and background separation, a BDT is used. This will be referred to, as before, as BDT_{ZZ}. BDT_{ZZ} is used to further

subdivide the channels, which defines the analysis categories. BDT_{ZZ} inputs remain unchanged from those used in the previous analysis; however, the BDT has been retrained using updated MC simulations. The training is done on the 4ℓ and $2\ell2\ell'$ cases separately, using ggF MC for signal (all mass points) and $qqZZ$ and $ggZZ$ background samples in the window $115 < m_{4\ell} < 130$ GeV. As in the previous analysis, events are divided into bins depending on their BDT score. To determine the best binning, the significance of a BDT bin is first defined. As the background is well understood, the asymptotic formula for the median expected significance is used, given by

$$z_i = \sqrt{2((S_i + B_i)\ln(1 + S_i/B_i) - S_i)}, \tag{6.9}$$

where S_i and B_i are the number of signal and background events in bin i respectively [53]. The total significance, z, is given by the quadrature sum over all bins. The value of z is investigated for several different numbers of bins. As a starting point, the values of z are first calculated for an even binning, then a hill-climbing algorithm is used to find optimal bin boundaries which maximise the value of z [54]. The values of z vs number of bins for the even and optimised cases are shown in Figure 6.10. As shown, there

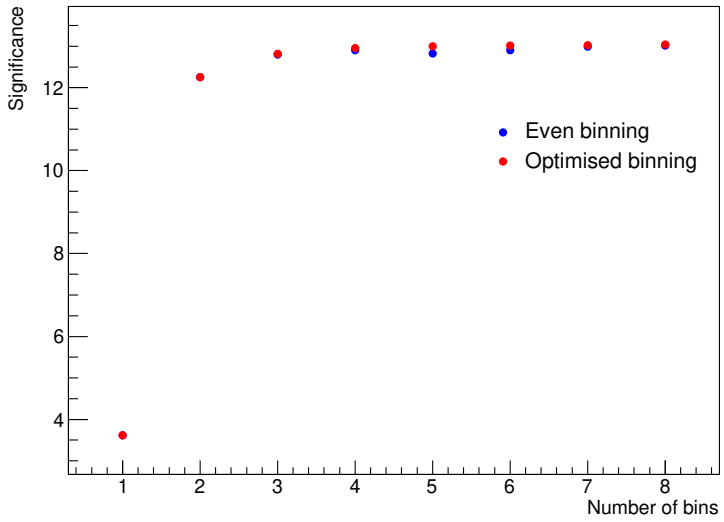
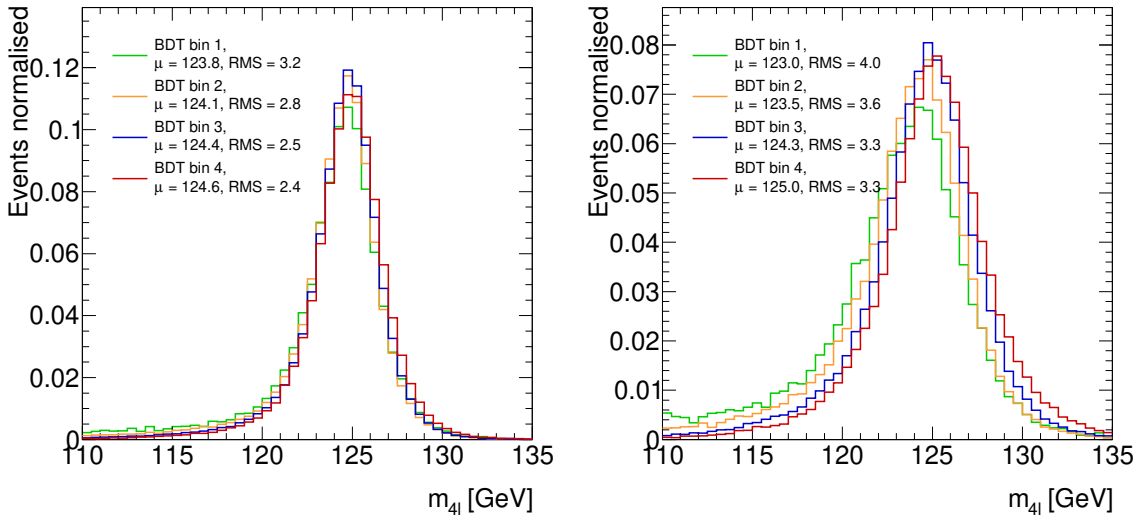


Fig. 6.10 Optimisation of the BDT binning, showing the values of bin significance, z vs number of bins for the even binning (blue) and optimised binning (red) cases.

is little improvement in z after reaching four bins. It is also clear from Figure 6.10 that the optimisation brings little benefit. For this reason, as in the previous analysis, four evenly-sized BDT_{ZZ} bins are used. The bins are referred to in increasing order of

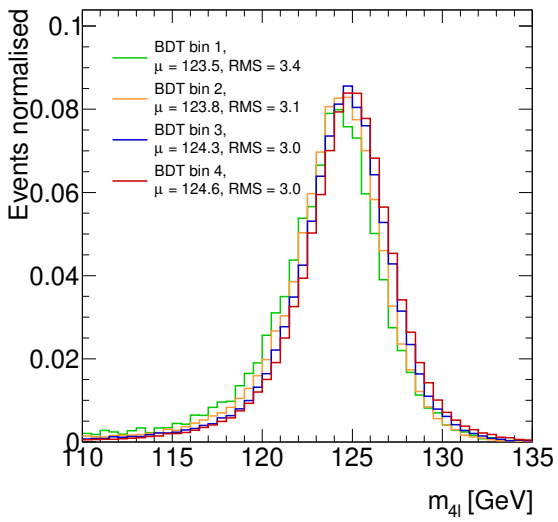
signal purity as BDT bins 1, 2, 3 and 4. The four BDT bins and the four channels define the 16 categories used in the analysis.

The distributions of $m_{4\ell}$ for the signal ggF production mode in each BDT bin are shown in Figure 6.11. This shows that events in the more signal-like BDT bins will tend

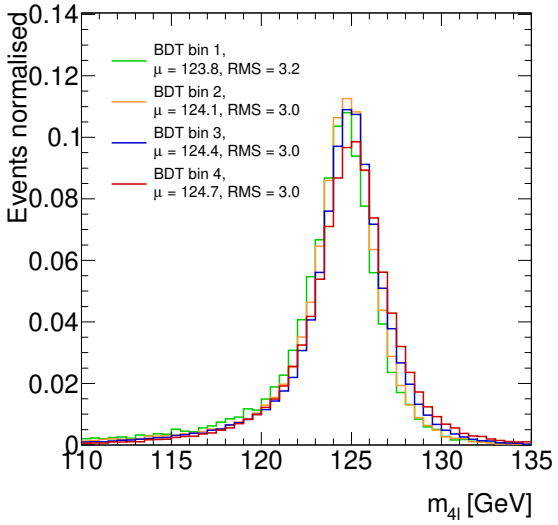


(a) 4μ

(b) $4e$



(c) $2\mu 2e$



(d) $2e2\mu$

Fig. 6.11 Distributions of $m_{4\ell}$ in each of the four BDT_{ZZ} bins in each channel for signal ggF. The mean of the distribution tends to be closer to m_{H} and the RMS lower for the more signal-like bins.

to have $m_{4\ell}$ closer to m_{H}. As the RMS for the more signal-like BDT bins is also lower, it is also the case that events with a higher BDT_{ZZ} score tend to be better-measured.

As the $m_{4\ell}$ distributions are a different shape in each of the categories, a model is built in each category independently. Once the 16 models have been built, they can be

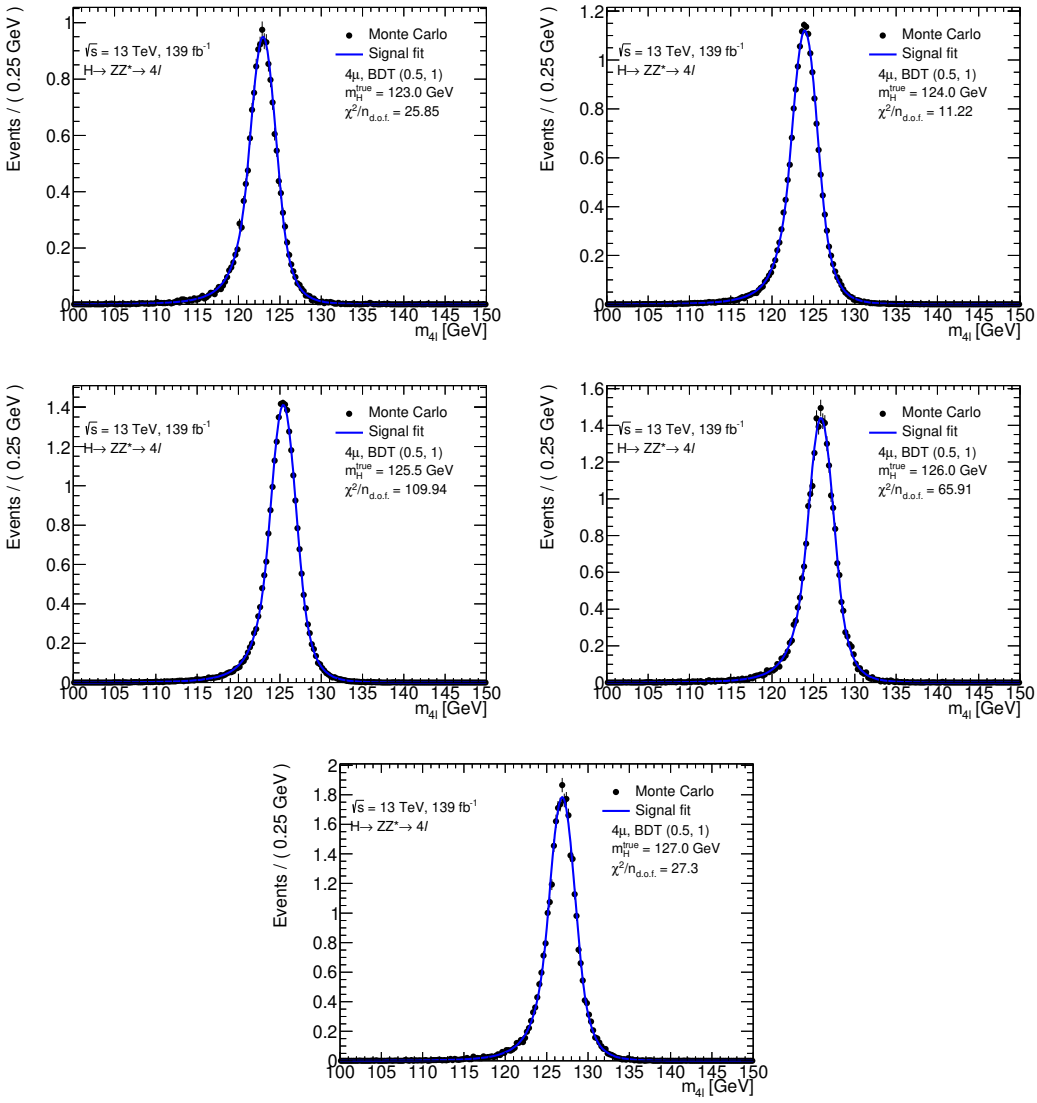


Fig. 6.12 Calibration plots for the analytic model in the 4μ channel, BDT bin 4.

fitted to the data simultaneously, sharing m_{H}. This model, which provides an analytic description of $m_{4\ell}$ in the 16 categories will be referred to as *the analytic model*.

To create the parametrisations for the analytic model, MC for the mass points 123, 124, 125.5, 126 and 127 GeV for the ggF and VBF production modes are used. Note that the 125 GeV mass point is not used here, as it is reserved for validation. One of these fits is performed per category, simultaneously across mass points with all parameters floating. The results of each of these fits in the 4μ BDT bin 4 category are

shown in Figure 6.12. Whilst the $\chi^2/n_{d.o.f.}$ values here are worse than in Figure 6.1, this does not have any significant impact on the mass measurement. The reason for the worsening is due to the additional constraint that the value of each tail parameter is the same across each mass point, so the tails do not fit as well causing a worsening of fit quality. As stated previously, the core of the distribution has the greater effect on the mass measurement, so this worsening of fit quality is not of concern.

A validation fit is then performed using the same procedure as before, m_{H} floating and all other parameters fixed, per category, to the 125 GeV mass point. Additionally, a simultaneous fit is also performed across all categories to produce a combined result. Figure 6.13 shows the results from these fits.

The combined measurement indicates closure in the fit, so the model is valid at the full MC statistics.

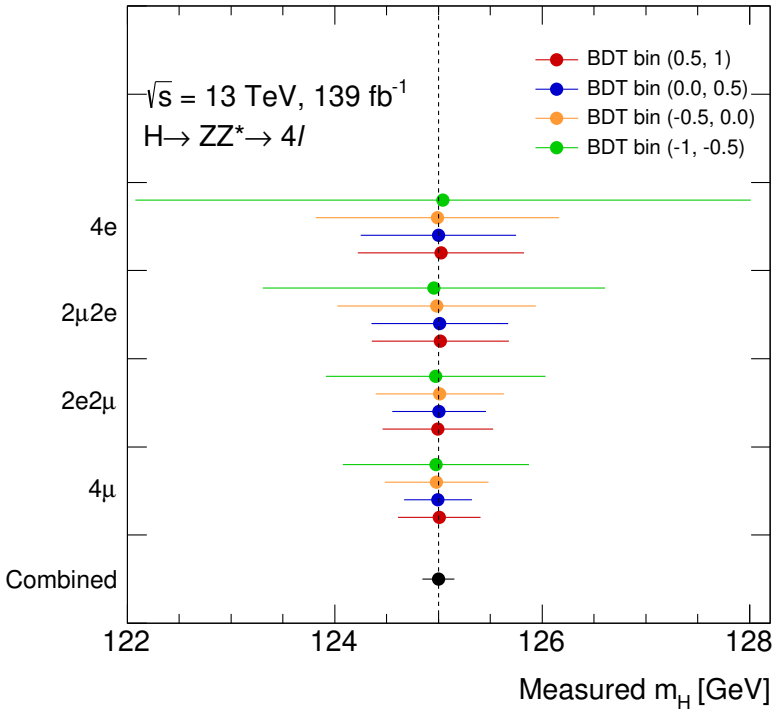


Fig. 6.13 Values of $m_{\mathrm{H}}^{\mathrm{measured}}$ from each category as well as one from a combined fit across all 16 categories. The fits show good closure. Uncertainty bands in this plot reflect the expected statistics, rather than the number of generated MC events.

Validation on expected statistics

Around 200 signal events are expected using the ATLAS full Run 2 dataset, so the model must also be tested with these statistics. Two methods are used to generate MC sets of this size. The first is to select a Poisson distributed number of events from the $H \to 4\ell$ ggF, VBF and VH MC samples for $m_{\mathrm{H}} = 125$ GeV with replacement. The mean number of events drawn here is that expected for the $H \to 4\ell$ process in each of the production modes at an integrated luminosity of 139 fb^{-1}. This will be referred to as *bootstrapped MC*. The second is to use the PDF derived from the calibration step, from which a Poisson-distributed number of events of mean expected at 139 fb^{-1}, are drawn. This is referred to as *toy data*. The validation procedure for both the toy data and bootstrapped MC is the same as in the full statistics case. This process of bootstrapping an MC sample or drawing toy data and performing a validation fit is repeated 10,000 to 100,000 times to generate a distribution of expected m_{H} values.

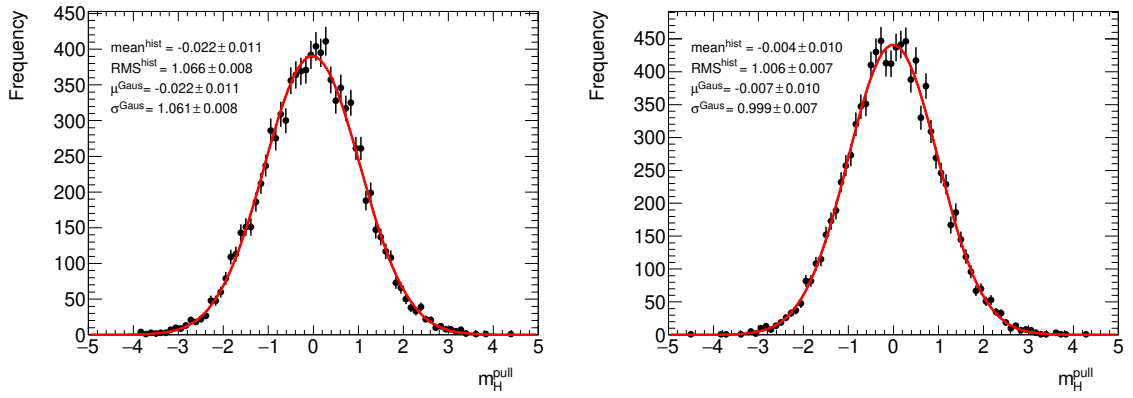
The bias and the goodness of the uncertainty calibration are determined through the use of *pull distributions* [55]. Suppose that a randomly distributed Gaussian variable x has a mean value of μ and a standard deviation of σ, the pull is defined as

$$p = \frac{x - \mu}{\sigma}. \tag{6.10}$$

Clearly, as x is Gaussian distributed, p will also be Gaussian distributed with a mean of 0 and a width of 1. The mean being displaced from 0 indicates a bias in the model and a width greater (less) than 1 indicates that errors have been underestimated (overestimated). The pull of m_{H} is defined as

$$m_{\mathrm{H}}^{\mathrm{pull}} = \frac{m_{\mathrm{H}}^{\mathrm{measured}} - m_{\mathrm{H}}^{\mathrm{true}}}{\sigma_{m_H}}, \tag{6.11}$$

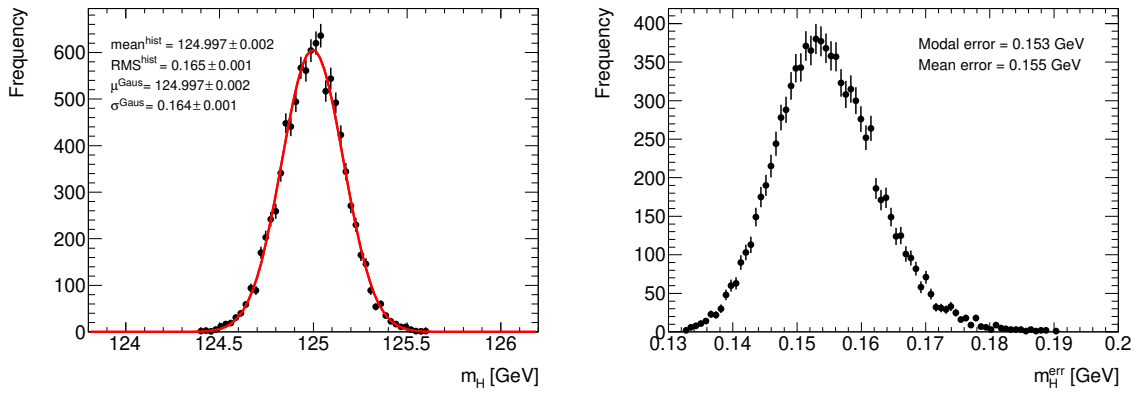
where σ_{m_H} is the uncertainty in $m_{\mathrm{H}}^{\mathrm{measured}}$, calculated from the covariance matrix from the fit. 100,000 bootstrapped MC datasets are created and the signal model fitted to them. From this, the value of $m_{\mathrm{H}}^{\mathrm{pull}}$ is extracted, the distribution of which is shown in Figure 6.14a. As the mean of the fitted Gaussian is less than 0, there is a small bias in the negative direction and, as the width is greater than 1, the uncertainties have been underestimated. To investigate the cause of the bias, toy datasets are generated from the model to check its self-consistency. The model is then fitted to these toy datasets using the same procedure as before and the resulting pull distribution is shown in Figure 6.14b. As the mean is consistent with 0 and the width with 1, this indicates that the bias is caused by the model not perfectly describing the MC. To determine



(a) (b)

Fig. 6.14 Pull distributions for $m_{\mathrm{H}}^{\mathrm{pull}}$ from fitting the model for $m_{\mathrm{H}}^{\mathrm{true}} = 125$ GeV to (a) bootstrapped MC and (b) toy data generated from the model. In both cases, all parameters other than m_{H} are fixed. In each case a Gaussian is fitted to the distribution. The distribution from toys shows no bias and good calibration of uncertainties and the bootstrapped shows a small negative bias and some underestimation of uncertainties.

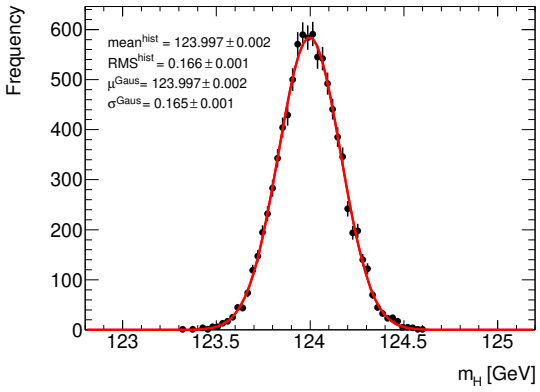
impact of this bias on the measurement of m_{H}, the distribution of $m_{\mathrm{H}}^{\mathrm{measured}}$ is also produced. Figure 6.15a shows the distribution of $m_{\mathrm{H}}^{\mathrm{measured}}$, which has a bias of 3 MeV. The expected uncertainty distribution is shown in Figure 6.15b, with the most probable



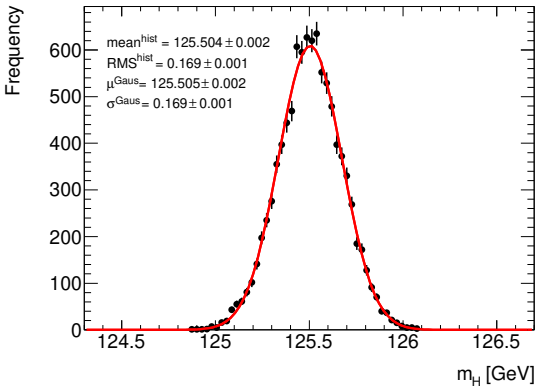
(a) (b)

Fig. 6.15 Results for bootstrapped MC using the ggF, VBF and VH production modes. Shown are (a) Distribution of $m_{\mathrm{H}}^{\mathrm{measured}}$ and (b) distribution of the uncertainties in $m_{\mathrm{H}}^{\mathrm{measured}}$ from fitting the model for $m_{\mathrm{H}}^{\mathrm{true}} = 125$ GeV 10,000 times to bootstrapped MC. All parameters other than m_{H} are fixed. A Gaussian fit is added to (a) with an expected mean of 125 GeV. The plot shows a negative bias in line with Figure 6.14. (b) shows that the most probable uncertainty in $m_{\mathrm{H}}^{\mathrm{measured}}$ is 153 MeV with the mean uncertainty being 155 MeV.

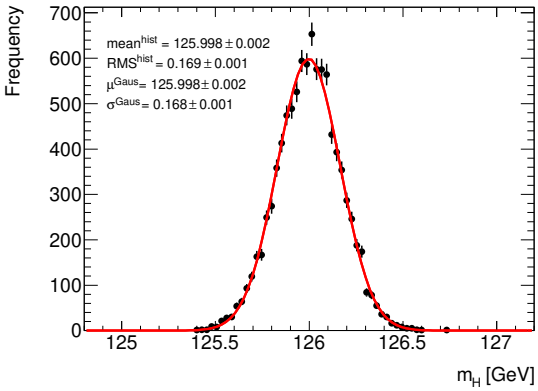
uncertainty in $m_{\mathrm{H}}^{\mathrm{measured}}$ being 153 MeV. To ensure the chosen parameterisation is valid for other mass points, models are built for different values of $m_{\mathrm{H}}^{\mathrm{true}}$. As before the "true" mass point to be used for validation is omitted in the calibration of each model, i.e. in the case where $m_{\mathrm{H}}^{\mathrm{true}} = 124$ GeV, only the 123, 125, 125.5, 126 and 127 GeV mass points are used[2]. The same procedure of validation is then used as before. The results from the the distributions of $m_{\mathrm{H}}^{\mathrm{measured}}$ for the mass points 124, 125.5 and 126 are shown in Figure 6.16. The distributions show the expected value of m_{H}, with biases of up to 5 MeV. The results of Figures 6.15 and 6.16 are summarised in Table 6.2.



(a) $m_{\mathrm{H}}^{\mathrm{true}} = 124$ GeV



(b) $m_{\mathrm{H}}^{\mathrm{true}} = 125.5$ GeV



(c) $m_{\mathrm{H}}^{\mathrm{true}} = 126$ GeV

Fig. 6.16 Distributions of $m_{\mathrm{H}}^{\mathrm{measured}}$ from validating the models for $m_{\mathrm{H}}^{\mathrm{true}} = 124$, 125.5 and 126 GeV on bootstrapped MC for their respective mass points. Each show a bias of up to 5 MeV.

[2]Note that in the case of $m_{\mathrm{H}}^{\mathrm{true}} = 125.5$ GeV, as no VH sample is available, those from $m_{\mathrm{H}}^{\mathrm{true}} = 125$ GeV are used, adjusting $m_{4\ell}$ and production cross sections to those expected at $m_{\mathrm{H}}^{\mathrm{true}} = 125.5$ GeV.

$m_{\mathrm{H}}^{\mathrm{true}}$ [GeV]	Mean of $m_{\mathrm{H}}^{\mathrm{measured}}$ [GeV]	Uncertainty in mean of $m_{\mathrm{H}}^{\mathrm{measured}}$ [GeV]	Bias in $m_{\mathrm{H}}^{\mathrm{measured}}$ [GeV]
124	123.997	0.002	0.003
125	124.997	0.002	0.003
125.5	125.505	0.002	0.005
126	125.998	0.002	0.002

Table 6.2 Summary of results of Figures 6.15 and 6.16. Means, uncertainties and biases of the distributions of $m_{\mathrm{H}}^{\mathrm{measured}}$ from fitting the model for various different values of $m_{\mathrm{H}}^{\mathrm{true}}$ 10,000 times.

Whilst some biases are seen in the validation at expected statistics, this is very small in comparison with the expected statistical uncertainty and can therefore be added as a systematic uncertainty.

6.2.2 Conditional parameterisation using per-event errors

The analytic model can be generalised to account for a per-event resolution. To make this change, the σ parameter for the DCB is replaced with a per-event resolution, σ_i, where i is the event number. The DCB thus becomes

$$P(m_{4\ell}|\sigma_i) = DCB(m_{4\ell}, \sigma_i; \mu, \alpha_{\mathrm{low}}, n_{\mathrm{low}}, \alpha_{\mathrm{high}}, n_{\mathrm{high}}), \tag{6.12}$$

and Equation 6.1 becomes

$$L(x_i) = \prod_i^N P(m_{4\ell}|\sigma_i). \tag{6.13}$$

To obtain per-event uncertainties on $m_{4\ell}$ (σ_i), the most obvious strategy would be to propagate uncertainties on the lepton energy and momentum. However, this tends to underestimate lepton uncertainty by around 10%-20%. In the previous analysis, this was solved by using Gaussian mixtures. As previously mentioned, this is very computationally intensive, especially given the increased number of events. The method employed is to estimate the σ_i for each event using a machine learning approach, namely a Quantile Regression Neural Network (QRNN) [56]. Most generally, a QRNN is used to predict the quantile of a target distribution. A QRNN is trained for each final state, on all the ggF MC mass points as training/testing sets using the Tensorflow library and Keras API [57, 58]. This targets the resolution on $m_{4\ell}$, $|m_{4\ell} - m_{4\ell}^{\mathrm{true}}|$, where $m_{4\ell}^{\mathrm{true}}$ is the true value of $m_{4\ell}$ (i.e. without detector smearing). A typical choice of target quantile is 68% as this corresponds to 1σ of a Gaussian distribution. However, as a

DCB is being fitted to the $m_{4\ell}$ distribution, the choice of the quantile which would have σ_i correspond to the width of the DCB is expected to be less than 68% due to the DCB's power law tails. At first, the 68% quantile is targeted. In order to determine the correct target quantile, the substitution $\sigma_i \to \text{SF} \cdot \sigma_i$ is made in Equation 6.12, where SF is a scale factor determined from fitting a DCB to ggF MC. Next, networks are trained targeting quantiles from 56% to 68% in 1% steps and the network for which SF is closest to 1 is chosen.

This model will be referred to as the per-event resolution or *PER model*. The PER model is built in the same 16 categories as previously defined and is fitted using an unbinned maximum likelihood fit, replacing the parameter σ with the product $\text{SF} \cdot \sigma_i$, where SF is a scale factor determined from the fit. These fits are shown in the 4μ BDT 4 category in Figure 6.17.

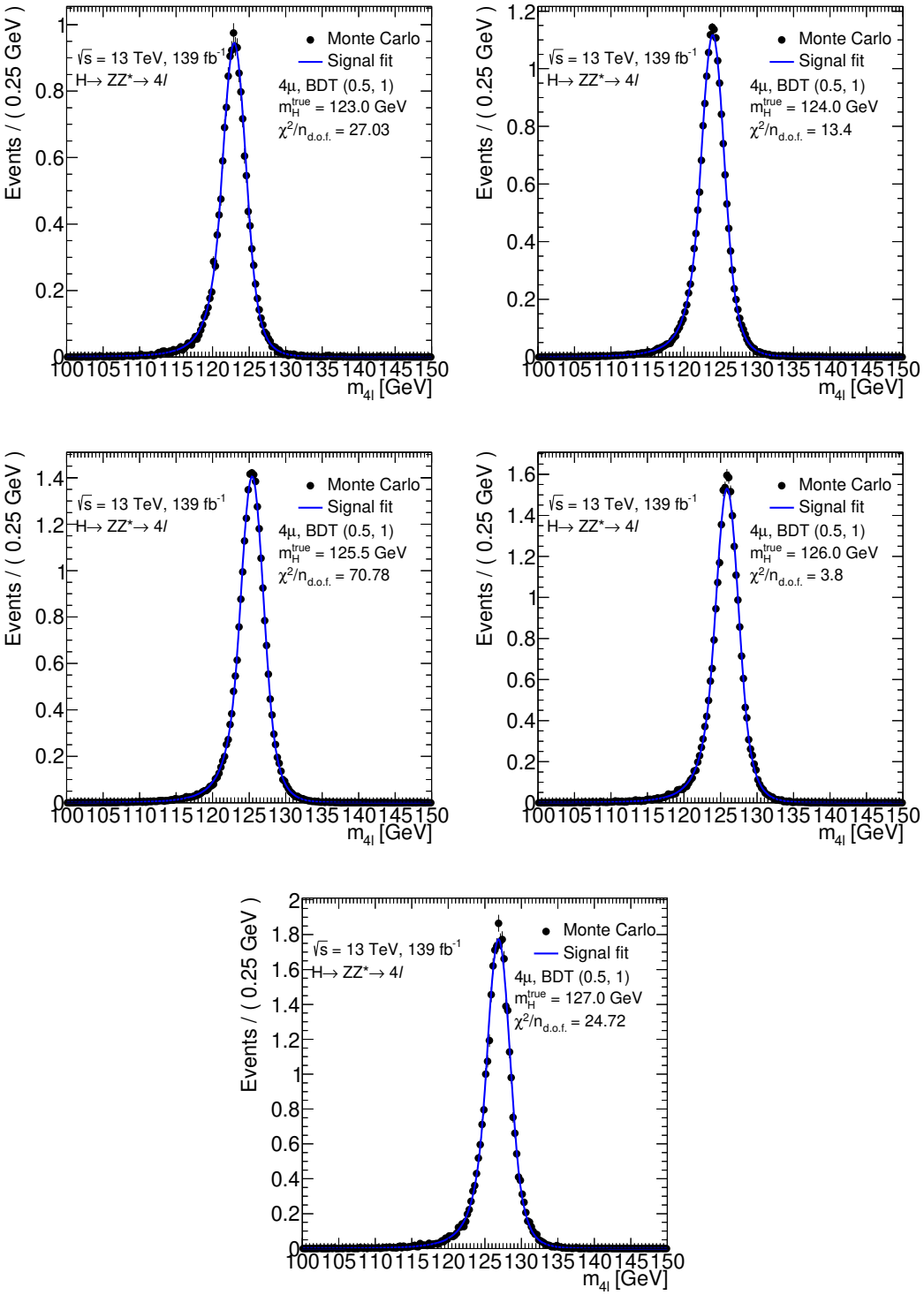


Fig. 6.17 Calibration plots for the PER model in the 4μ channel, BDT bin 4.

6.3 Background model

In addition to the $qqZZ$ and $ggZZ$ components to the irreducible ZZ^* background used in the previous analysis, an electroweak ZZ^* component is added (EW ZZ). Previously, $t\bar{t}+Z$ and VVV production were considered for the sub-leading irreducible backgrounds. This is updated to include several more backgrounds from single top and $t\bar{t}$ in association with other objects. These are collectively referred to as tXX. A list of all background MC samples is shown in Table 6.3. It should be noted that, from the N_{expt} column, most of the contributions to the background are negligible.

6.3.1 Background modelling

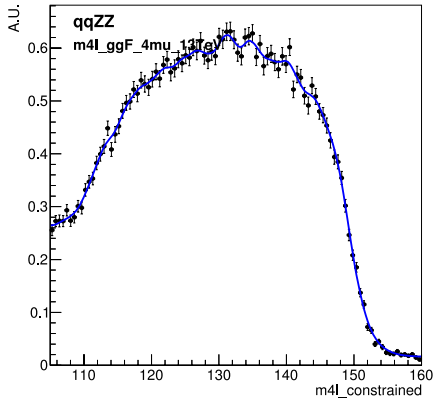
As in the previous analysis, background shapes are modelled through the smoothing of the MC using a method of adaptive KDE smoothing with Gaussian kernels [36]. The smoothing is configured through the use of the smoothing parameter, ρ. Smaller values of ρ conserve more features but are more susceptible to statistical fluctuations, whereas larger values of ρ tend to smooth over fluctuations better but do not preserve features of the distribution as well. A demonstration of the effect of varying ρ is shown in Figure 6.18 for three values of ρ. The values of ρ are chosen for each category and each process individually, with the best value being determined by eye. In order to enhance the MC statistics in the region around the Higgs peak, a $qqZZ$ MC sample with greater statistics, filtered in the range $100 < m_{4\ell} < 150$ GeV, is used. As the range of the fit extends in the positive direction beyond the coverage of the filtered sample, an additional sample is used to cover this region. This sample, however, does not have as high MC statistics. This large difference in statistics can lead to the smoothed shape not being a good description of the MC in the region where the two samples overlap. To solve this problem, two approaches to smoothing were investigated. The first is to smooth each sample individually and sum the two smoothed shapes. The second is to combine the MC samples and produce a single smoothed shape. A comparison of these two methods is shown in Figure 6.19. The approach of smoothing the components individually (red) tends to produce the more accurate description, so this approach is used.

6.4 Signal and background normalisation

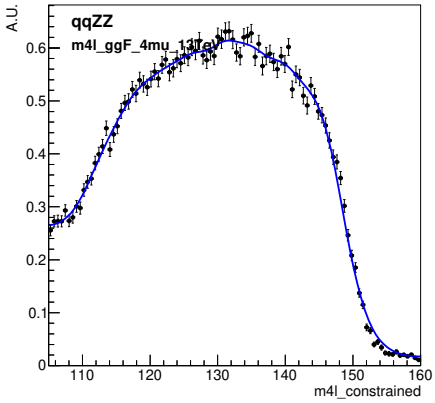
As the $m_{4\ell}$ range has been expanded for this analysis, the sidebands, which have relatively few signal events, can be used to obtain the background normalisation from

Process	$m_{4\ell}$ range filter [GeV]	N_{MC}	N_{expt}
$qqZZ$	100-150	2001443	254.9
	Inclusive	2357387	52.0
	High	1885864	0.42
$ggZZ$	0-130	10009	2.95
	Inclusive	366184	11.2
EW ZZ	None	107716	1.31
$t\bar{t}\mu\mu$	100-150	226494	3.32
	Inclusive	25821	0.95
$t\bar{t}ee$	100-150	179057	2.61
	Inclusive	21539	0.67
$t\bar{t}$W	None	1020	0.09
	None	106	0.03
$t\bar{t}WW$	None	3173	0.006
	None	19	0.04
$t\bar{t}WZ$	None	14157	0.02
$t\bar{t}ZZ$	None	42589	0.02
$t\bar{t}\gamma$	None	46	0.86
$t\bar{t}Z\gamma$	None	11850	0.02
$tllq$	None	6197	0.22
	None	1908	0.07
tWZ	None	1692	0.89
	None	42410	0.002
$tt\bar{t}$	None	41	0
	None	1193	0.004
$t\bar{t}t\bar{t}$	None	1061	0.03
	None	59	0.003
$WWZ \to 4\ell 2\nu$	None	6869	0.76
$WZZ \to 5\ell\nu$	None	38699	0.15
$ZZZ \to 6\ell$	None	54851	0.03
$ZZZ \to 4\ell 2\nu$	None	26228	0.08

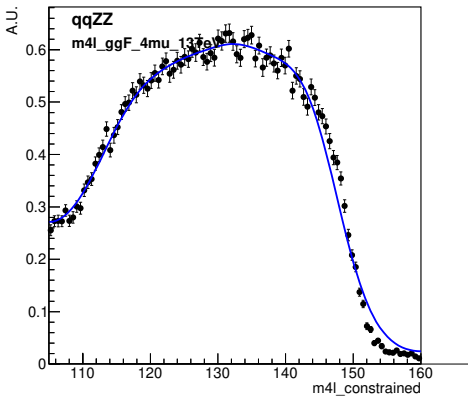
Table 6.3 A list of all the simulated background processes. For the backgrounds $qqZZ$, $ggZZ$, t$\bar{\text{t}}\mu\mu$ and t$\bar{\text{t}}ee$, multiple samples are used with a filter on $m_{4\ell}$ applied. This is so MC samples with enhanced statistics in a desired range can be generated. Overlap between such samples is accounted for by the weighting of events. N_{MC} is the number of unweighted MC events which pass selection and N_{expt} is the expected number of events at 139 fb^{-1} in the $105 < m_{4\ell} < 160$ GeV range.



(a) $\rho = 1.0$



(b) $\rho = 2.0$



(c) $\rho = 3.4$

Fig. 6.18 Demonstration of the effect of varying the smoothing parameter, ρ, on the $m_{4\ell}$ distribution in the 4μ channel for the $qqZZ$ background, shown for values of $\rho = 1, 2, 3.4$. In the case of $\rho = 1$ (a), nearly all features are preserved but the shape is susceptible to fluctuations. In the case of $\rho = 3.4$ (c), these fluctuations are smoothed over, but some of the shape is lost in doing so. Note that the sharp drop around $m_{4\ell} \approx 150$ GeV is not physical, and is caused by the filtering of events.

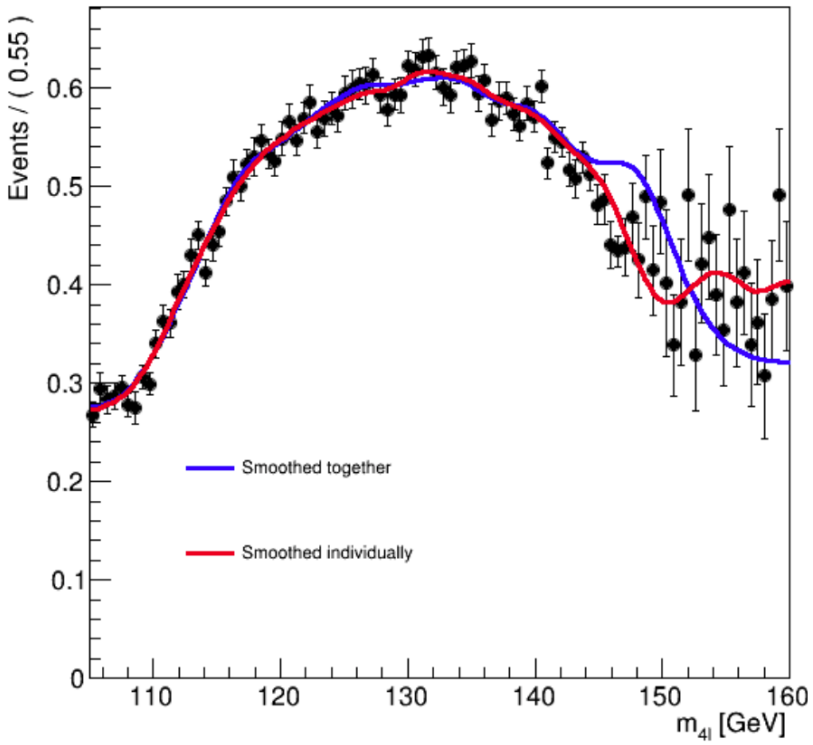


Fig. 6.19 Comparison of different techniques to smooth the two MC samples for $qqZZ$ in the 4μ channel. The first, filtered, sample covers the range 100-150 providing greater MC statistics here. The second, the inclusive sample, covers all other parts of the $m_{4\ell}$ spectrum. The first approach is to smooth each component individually (red) and the second is to smooth the components together (blue).

data. The expected yields of $qqZZ$ and $ggZZ$ are estimated from MC and the expected number of background events per category is denoted $N_{ij}^{\text{exp, bkg}}$, where i and j are the channel and BDT bin respectively. One normalisation is used per BDT bin, so the number of background events in a BDT bin j is parametrised as

$$N_{ij}^{\text{obs, bkg}} = \nu_j^{\text{bkg}} \cdot N_{ij}^{\text{exp, bkg}}, \tag{6.14}$$

where $N_{ij}^{\text{obs, bkg}}$ is the observed number of background events in a category ij and ν_j^{bkg} is the relative normalisation for BDT bin j, determined from the fit to data.

The yields for the EW ZZ, tXX and VVV processes are estimated from MC and their relative normalisations are fixed to 1. The yields of the reducible backgrounds are determined from dedicated data control regions [38] and their relative normalisations are also fixed to 1.

For the signal, previously the expected number of signal events was determined from branching ratio, cross section, luminosity and detector acceptance. As with the irreducible background, this is replaced with expected events being determined from MC, using the mass point $m_{\mathrm{H}} = 125$ GeV. The signal is normalised in the same manner as the ZZ^* background: that is, the yield for all signal processes is determined from MC and their relative normalisations in BDT bin j, ν_j^{sig}, are given by

$$N_{ij}^{\mathrm{obs,\ sig}} = \nu_j^{\mathrm{sig}} \cdot N_{ij}^{\mathrm{exp,\ sig}}, \tag{6.15}$$

where $N_{ij}^{\mathrm{obs,\ sig}}$ is the observed number of signal events in a category ij and $N_{ij}^{\mathrm{exp,\ sig}}$ is determined from MC.

6.5 Systematic uncertainties

The systematic uncertainties considered broadly fall into two categories: normalisation, which affects the number of expected events, and shape, which affects the shape of the $m_{4\ell}$ distribution. To implement these into the model, each MC sample is modified to reflect the up and down variations for each systematic. The approach to implementing these variations in the case of normalisation systematics is to adjust the weighting of events in the MC simulations to reflect the up and down variations. In the case of shape systematics, two new sets of MC samples are generated for both the up and down variations for each systematic for each process simulated. In both of these cases, the up and down simulations are used to derive two new models for that systematic uncertainty. These up and down models as well as the nominal model are interpolated with a normalised *nuisance parameter*, θ.

In the case of the signal model, a new DCB is fitted to each of the up and down MC samples for each systematic. For this fit, only the b_μ and SF parameters (or σ for the model without per-event resolution) are allowed to float and all other parameters are fixed to their nominal values. The systematic variation on model parameter, a is related to the normalised nuisance parameter, θ, by the relation

$$\theta(a) = \frac{a - a_{\mathrm{nom}}}{\left|a_{\mathrm{up/down}} - a_{\mathrm{nom}}\right|}, \tag{6.16}$$

where a_{nom} is the nominal value of a and $a_{\mathrm{up/down}}$ are the up and down variations of the parameter a. As the value of a_{up} is not necessarily the same as the value of a_{down},

in practice, $\theta(a)$ is a piecewise function containing two sub functions for the cases of $a > a_{\text{nom}}$ and $a < a_{\text{nom}}$.

Thus, a value of $\theta = 0$ corresponds to the nominal model and a value $\theta = 1$ would correspond to the up model. The nuisance parameter is added to the likelihood as a Gaussian constraint, so the likelihood for a PDF with a single parameter a becomes

$$L(x_i) = \prod_i^n P(x_i; a) G(\theta(a); 0, 1), \tag{6.17}$$

where G is a Gaussian PDF of mean 0 and width 1. To generalise this process for an arbitrary number of systematics, the value of the parameter a will depend on a vector of nuisance parameters, $\boldsymbol{\theta}$, via some continuous function $f(\boldsymbol{\theta})$. To determine the parameters of f, the values for each nuisance parameter, whilst keeping all others at 0, are determined using the method above. To determine the value of a for all values of $\boldsymbol{\theta}$, a method of piecewise-linear interpolation is used. For each systematic, another Gaussian is added to the likelihood which depends solely on the nuisance parameter corresponding to that systematic.

For the background model, smoothed templates are created for each up and down variation and are interpolated using moment morphing [59], with the morphing controlled by the parameters $\boldsymbol{\theta}$.

In total, 74 sources of shape systematic relating to electron energy scale and resolution and 5 from muon momentum scale and resolution are considered. Distributions of $m_{4\ell}$ for the signal for the dominant electron systematic, the electron energy scale, are shown in Figure 6.20. This systematic tends to cause a shift in $m_{4\ell}$ which increases the mean of the distribution by approximately 50 MeV in the $4e$ case. The b_μ parameter of the model in the $4e$ channel is shifted by a similar amount, whilst the change to SF/σ is relatively small. Similarly, the dominant systematic affecting muons, the muon momentum scale, tends to cause a shift to the $m_{4\ell}$ distribution. As Figure 6.21 shows, the shift in the 4μ channel is $\sim$ 100 MeV and the width changes by an amount $<$ 10 MeV. As muons tend to have a better $m_{4\ell}$ resolution than electrons, this is the dominant systematic in the measurement.

Normalisation systematics accounted for include the relative amounts of ggF and VBF signal events. As the kinematics of a VBF event differ slightly from those of a ggF event, inaccurate estimation of the relative amounts of the two processes can lead to a bias. The uncertainties on the cross sections of ggF and VBF are used to estimate this. The effect on the value of σ is approximately 1 MeV, while the effect on b_μ is $<$ 1 MeV.

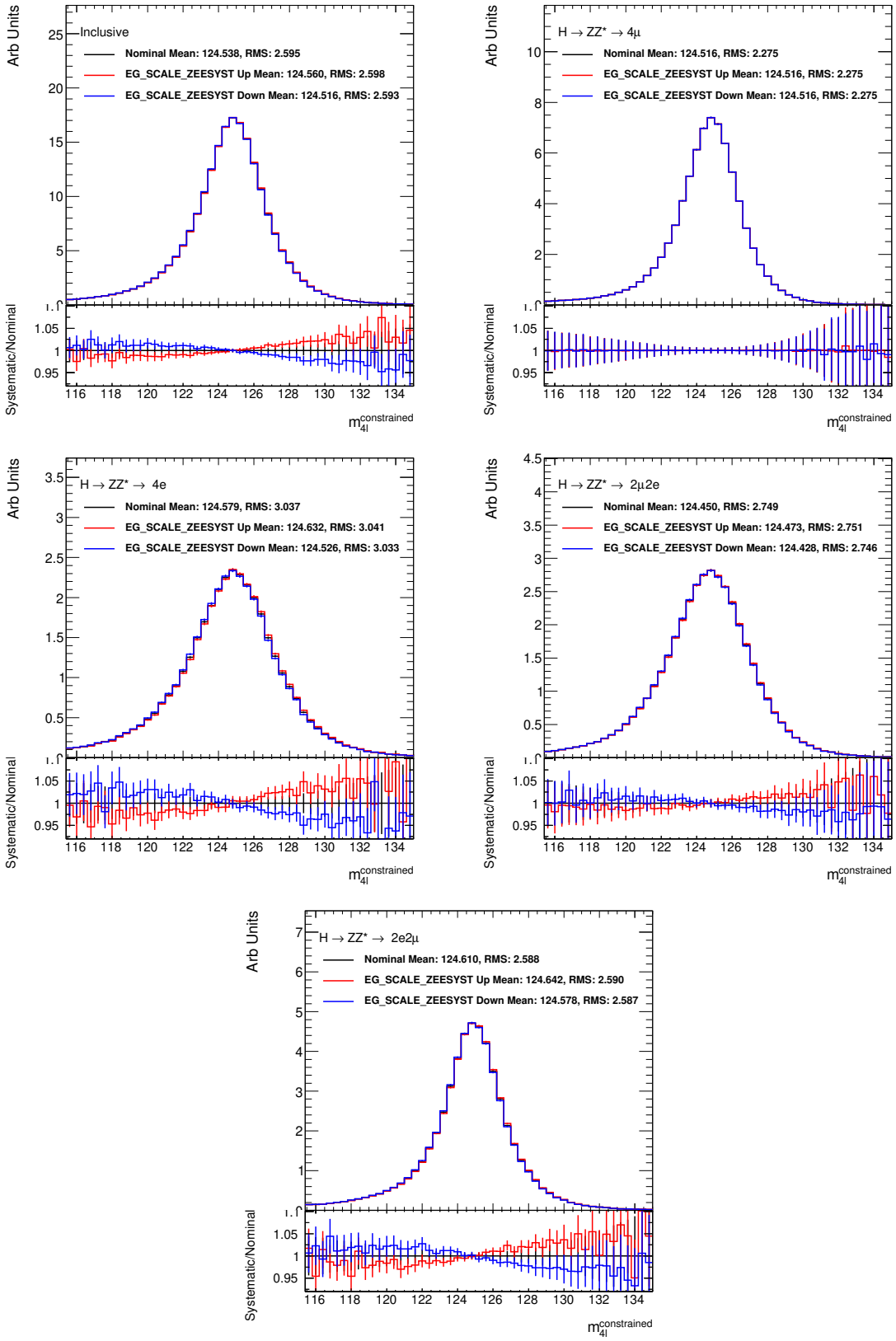


Fig. 6.20 Systematic variations of $m_{4\ell}$ distributions for the signal model of the electron energy scale systematic in each channel and inclusive of all channels.

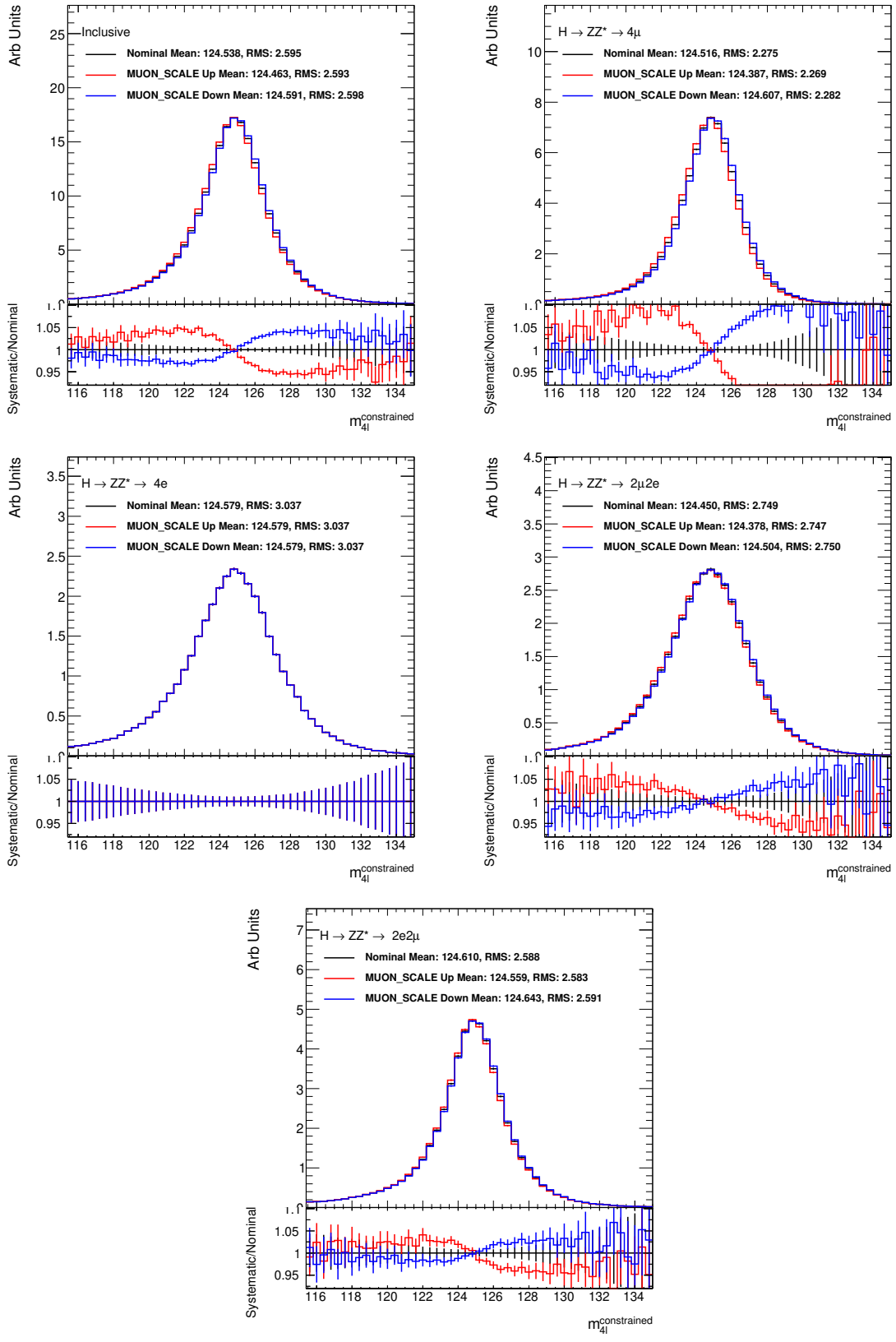


Fig. 6.21 Systematic variations of $m_{4\ell}$ distributions for the signal model of the muon momentum scale systematic in each channel and inclusive of all channels.

Other systematics considered include two biases. The larger of the two is the bias found when fitting to bootstrapped MC. This causes a bias in the measurement of 3 MeV (see Table 6.2).

The second is the bias due to the estimator. To investigate this bias, a DCB is created with the parameters $\mu = 0$, $\sigma = 1$, $\alpha_{\text{low}} = \alpha_{\text{high}} = 1$, $n_{\text{low}} = 1.2$ and $n_{\text{high}} = 10$. These are chosen as they are similar to the values expected for the 4μ channel. From this simple model, a toy dataset is generated with a Poisson distributed number of events of mean 100. The model is then fitted back to the dataset with μ floating and all other parameters fixed to their original values. The pull of μ is calculated and the procedure of generating a toy dataset and fitting to this is repeated 10^7 times. The distribution of the pull is then plotted and compared to the expected distribution (Gaussian with mean of 0 and width of 1). The distribution is shown in Figure 6.22. The mean of the fitted Gaussian is -0.0081 ± 0.0003, i.e. 25 standard deviations

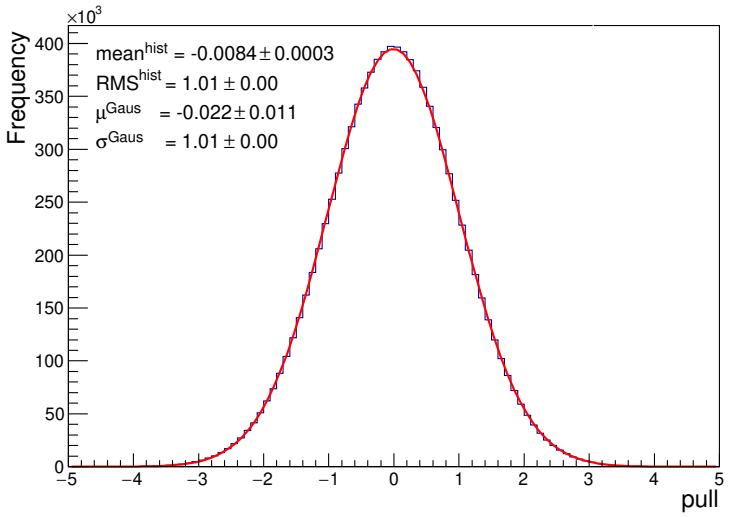


Fig. 6.22 Pull of μ from fitting a DCB to a toy dataset generated from that same DCB with μ floating and all other parameters fixed. A Gaussian fit to the distribution is also shown. The parameters of this Gaussian show a bias in the negative direction.

from the expected mean of 0. It should however be noted that, in general, maximum likelihood estimators are biased with a bias proportional $1/n_{evts}$ [47], where n_{evts} is the number of events. To check that this is indeed the case, the bias is studied as a function of number of events to see if this pattern holds true. The same procedure as detailed above is repeated with various numbers of toy events between 50 and 1000. The variation in bias with number of events is shown in Figure 6.23 along with a fit of the form $a + b/n_{evts}$ (with both a and b measured in GeV). If the distribution behaves

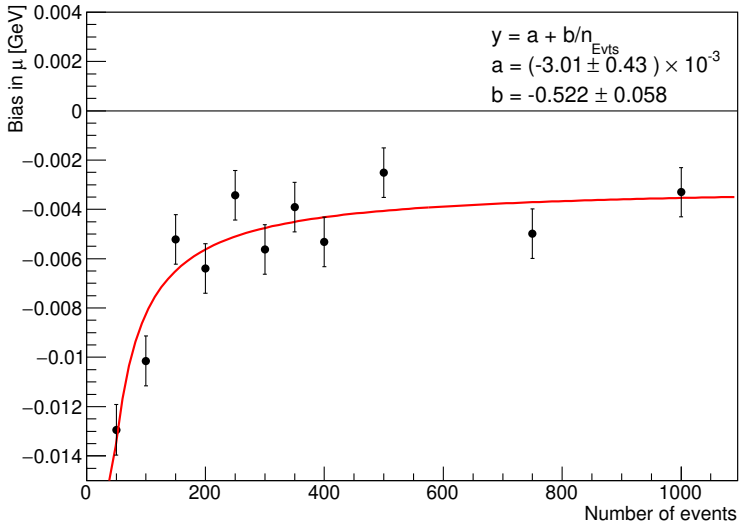


Fig. 6.23 Mean of the fitted Gaussian to the distribution of pulls on μ from fitting a DCB to toy data with μ floating, against number of events. The bias is expected to be proportional to $1/n_{evts}$, so a function of the form $a + b/n_{evts}$ is fitted, with the expectation that $a = 0$, where both a and b are measured in GeV. It is found that $a < 0$, so the bias is not proportional to $1/n_{evts}$.

as expected, then $a = 0$. However, from the fit $a = (-3.01 \pm 0.43) \times 10^{-3}$, which is not consistent with expectation. This indicates that the estimator bias decreases more slowly than expected. To check that this is due to the estimator rather than the implementation of the function, the same test is performed using a CB with parameters $\mu = 0$, $\sigma = 1$, $\alpha = 1$ and $n = 1.2$. The test of fitting to a toy dataset of 200 events using a CB with 10^6 iterations returns a mean of -0.0240 ± 0.0010. The bias in fitting a CB to toy data is also investigated when different parameters are floating/fixed, shown in Figure 6.24. Having σ floating moves the bias to the positive direction. Additionally setting α free increases the bias further. A fit of the form $a \times n_{evts}^{b}$ is also shown. As $b \approx 0.5$ This indicates that the bias rather than being $\propto 1/n_{evts}$ is closer to $\propto 1/\sqrt{n_{evts}}$.

To estimate the size of the expected bias in the measurement of m_{H}, a similar test to the above is performed using the calibrated analytic DCB model. Toy datasets are generated from the model in each of the 16 categories and the model is fit to these, with m_{H} as the sole free parameter. This process is repeated 10,000 times. The distribution of measured m_{H} values from this is shown in Figure 6.25 with a Gaussian fit applied. As the mean of the fitted Gaussian is 124.999 ± 0.002 (the true value being 125), the estimator bias is taken to be 1 ± 2 MeV.

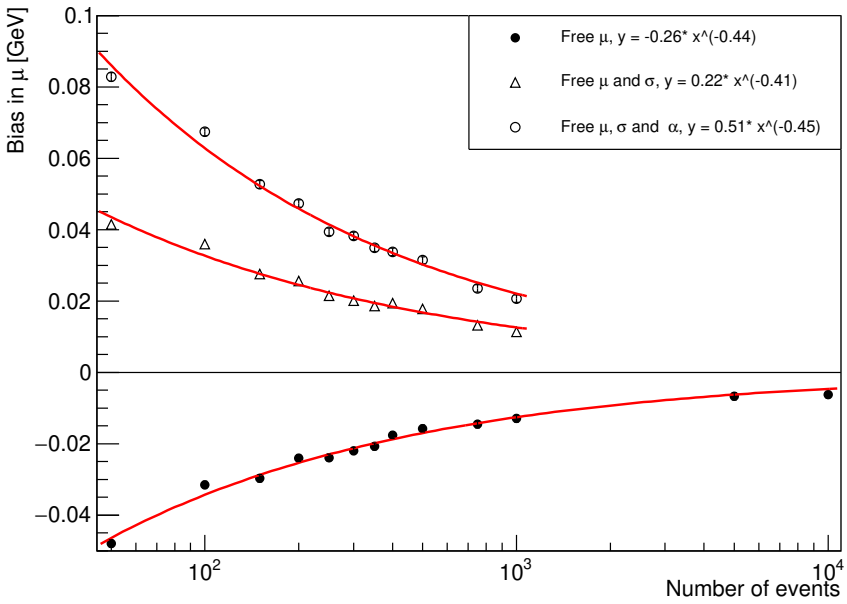


Fig. 6.24 Bias from pulls of μ of a CB fit to toy data generated from the same CB, against number of events. The different sets of points correspond to different parameters of the CB floating/fixed. The closed circles are for the case of only μ floating (with all parameters fixed), open triangles for μ and σ floating and open circles for μ, σ and α floating. A fit of the form $a \times n_{evts}^{b}$ is also shown.

From these studies, it is expected that the muon momentum scale will be the most important systematic uncertainty with the electron energy scale the second most important. Smaller contributions are expected to be seen from the electron and muon resolution and other systematics also having a very small impact.

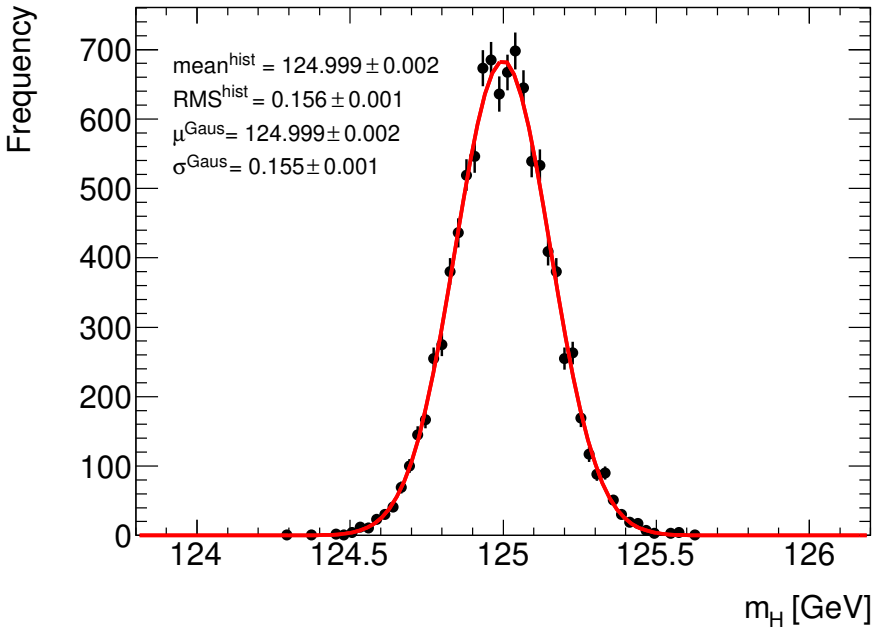


Fig. 6.25 Distribution of m_{H} values from fitting the calibrated analytic model to toy data generated from the model. This shows that the bias due to the estimator is 0 within uncertainty.

Chapter 7

Results of the Higgs boson mass measurement at 139 fb^{-1}

The measurement of m_{H} is performed using a profile likelihood ratio maximisation [60, 53]. The profile likelihood is defined as

$$\lambda(m_{\mathrm{H}}) = \frac{L(m_{\mathrm{H}}, \hat{\hat{\boldsymbol{\theta}}}(m_{\mathrm{H}}))}{L(\hat{m_{\mathrm{H}}}, \hat{\boldsymbol{\theta}})}, \tag{7.1}$$

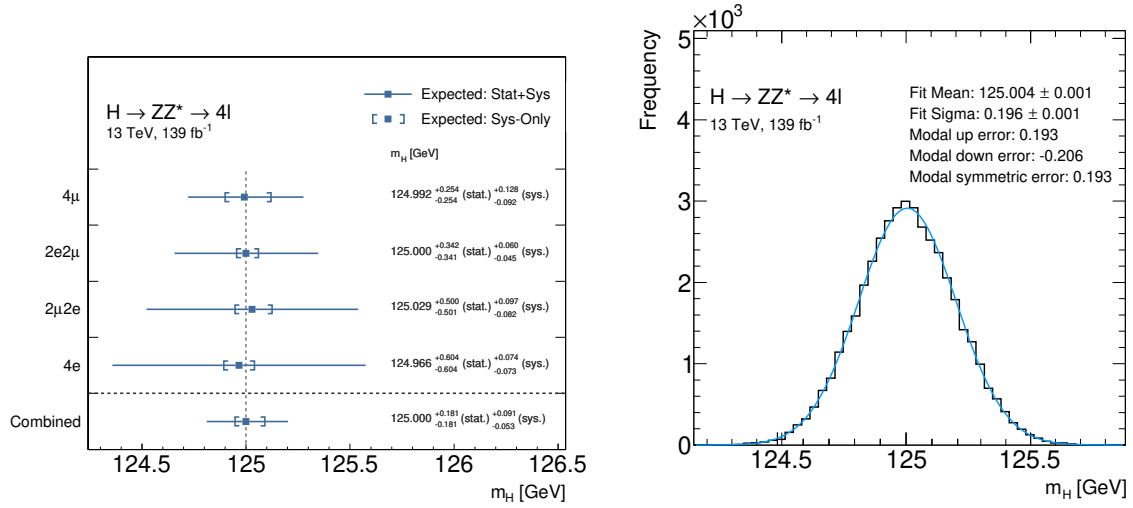
where the denominator is the unconditional maximised likelihood, i.e. $\hat{\boldsymbol{\theta}}$ and $\hat{m_{\mathrm{H}}}$ are the maximum likelihood estimates of m_{H} and nuisance parameters $\boldsymbol{\theta}$. In the numerator is the conditional maximised likelihood; here $\hat{\hat{\boldsymbol{\theta}}}$ are the values of $\boldsymbol{\theta}$ which maximise L for a given m_{H}. This is transformed, for convenience, to the negative log profile likelihood ratio, $-2\ln\lambda$. The best estimate of m_{H} is that which minimises $-2\ln\lambda$. These fits are performed with m_{H}, the four signal relative normalisations, ν_j^{sig}, the four ZZ^* background relative normalisations, ν_j^{bkg}, and all nuisance parameters floating. The resulting uncertainty on m_{H} will correspond to the total uncertainty. To determine only the statistical uncertainty, the same fit is performed, but with all nuisance parameters, $\boldsymbol{\theta}$, fixed to 0.

7.1 Expected results

7.1.1 PER model

Two tests on the PER model are performed, one by fitting to the full MC statistics and the second by fitting to bootstrapped MC. For the fitting to the full MC statistics, five

such fits are performed. The first four are combined fits per channel, simultaneously across each of the four BDT bins, with the other being a combined fit across all 16 categories. The results from this are shown in Figure 7.1a. The fit to bootstrapped MC is performed in the same manner as that performed on the signal model. The resulting distribution of m_{H} is shown in Figure 7.1b along with a Gaussian fit. The



(a) (b)

Fig. 7.1 Expected results for the PER model (a) results from fits to the full MC statistics and (b) distribution of m_{H} from fitting to bootstrapped MC with a Gaussian fit.

model closes with an expected total uncertainty of 196 MeV in the fit to full MC and 193 MeV in fits to bootstrapped MC. Shown in Figure 7.2 is the expected systematic ranking for the PER model. As in the previous analysis, muon momentum scale and electron energy scale uncertainties are the dominant systematics. None is significantly pulled or constrained.

7.1.2 Model without per-event resolution

The same procedure described above is performed for the model without per-event resolution. From these tests, the model closes with an expected total uncertainty of 199 MeV in the fit to full MC and 196 MeV in fits to bootstrapped MC. Shown in Figure 7.4 is the expected systematic ranking for the model without accounting for per-event resolution. As expected, the muon momentum scale and the electron energy scale are the dominant systematics. No systematics are significantly pulled with respect

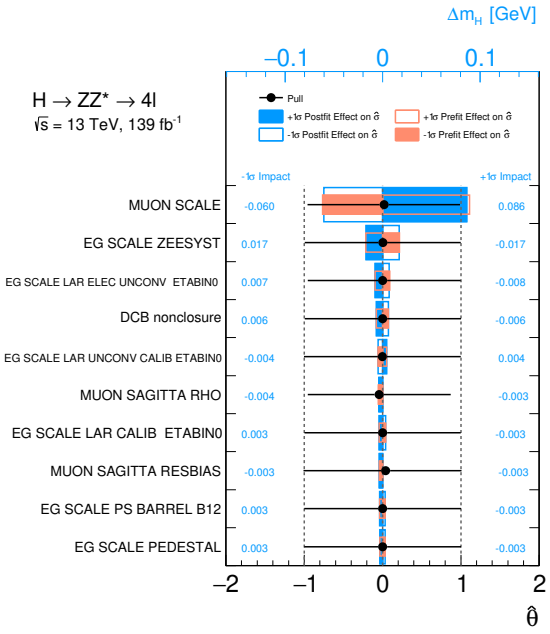
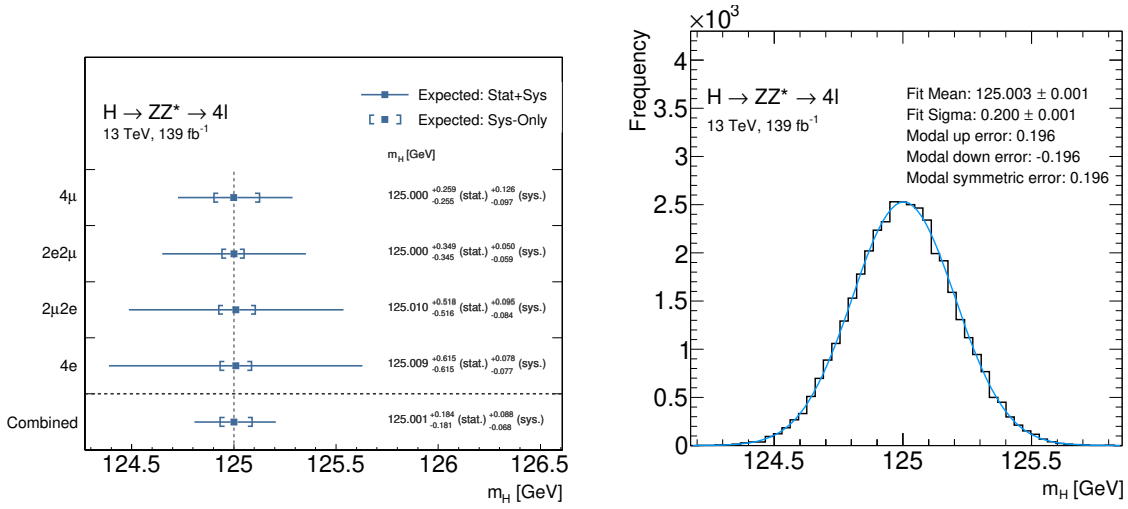


Fig. 7.2 Expected systematic ranking for the PER model. The red and blue bars indicate the prefit and postfit effects of a particular systematic on the uncertainty respectively (top x-axis). The black points indicate the amount by which the corresponding nuisance parameter is pulled from its nominal value (bottom x-axis).



(a) (b)

Fig. 7.3 Expected results for the model without accounting for per-event resolution (a) results from fits to the full MC statistics and (b) distribution of m_{H} from fitting to bootstrapped MC with a Gaussian fit.

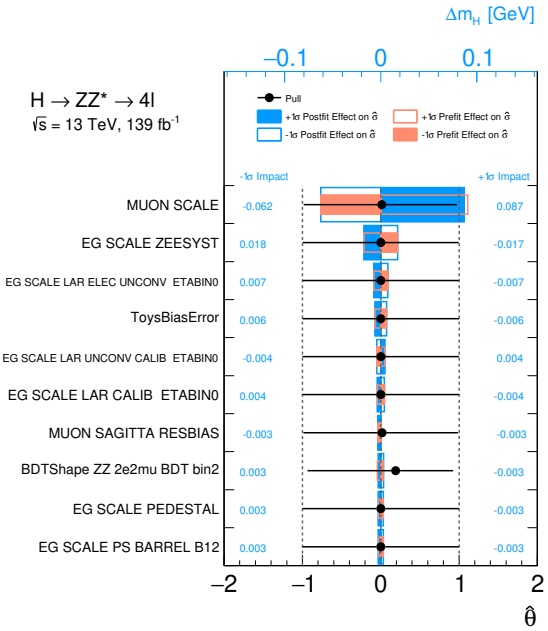


Fig. 7.4 Expected systematic ranking for the model without accounting for per-event resolution.

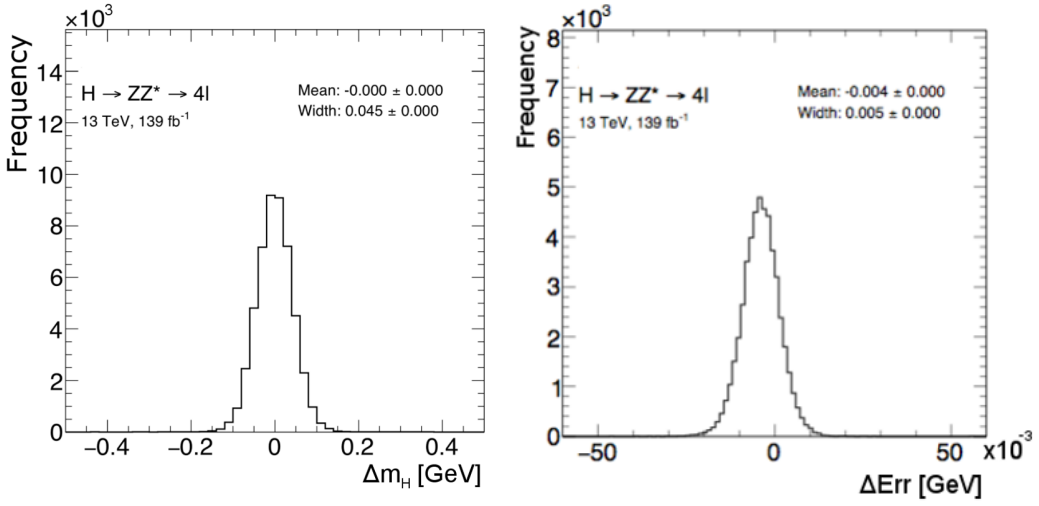
to their nominal value with the exception of the BDT shape uncertainty in the $2e2\mu$ BDT bin 2 category.

7.1.3 Compatibility between models

In order to verify that the addition of per-event uncertainties does not bias the model, it is fitted twice to the same sets of bootstrapped MC with and without accounting for per-event resolution. The difference in the value of m_{H} returned is investigated. The results from this are shown in Figure 7.5a. Accounting for per-event resolution, on average, returns the same result as when not accounting for it with a spread of 45 MeV, so the addition of per-event resolution does not bias the model. Shown in Figure 7.5b is the distribution in differences in uncertainty between the two cases when fitting to the same set of bootstrapped MC. On average, accounting for per-event resolution improves the uncertainty on m_{H} by 4 MeV, with a spread of 5 MeV.

7.2 Observed results

Table 7.1 shows the expected and observed yields for each of the channels around the Higgs boson peak, $115 < m_{4\ell} < 130$ GeV. Slightly more $2e2\mu$ and slightly fewer $4e$



(a) (b)

Fig. 7.5 Compatibility between the model with and without per-event resolution. (a) shows the difference in measured m_{H} between when fitting to the same sets of bootstrapped MC. As the peak is centred on 0, no bias is observed with a standard deviation of 45 MeV. (b) shows the difference in uncertainty between the same sets of bootstrapped MC. On average, the accounting for per-event resolution improves uncertainty estimation by 4 MeV with a standard deviation of 5 MeV.

Channel	Higgs ($m_{\mathrm{H}} = 125$ GeV)	ZZ^*	Other backgrounds	Total expected	Observed
4μ	78 ± 5	38.0 ± 2.1	2.85 ± 0.18	119 ± 5	115
$4e$	35.3 ± 2.6	15.0 ± 1.5	2.91 ± 0.33	53.2 ± 3.1	42
$2e2\mu$	53.0 ± 3.1	26.1 ± 1.4	2.98 ± 0.19	82.0 ± 3.4	96
$2\mu2e$	40.1 ± 2.9	17.3 ± 1.3	3.6 ± 0.5	61.0 ± 3.2	57
Inclusive	206 ± 13	96 ± 6	12.2 ± 1.0	315 ± 14	310

Table 7.1 Expected and observed yields for events within $115 < m_{4\ell} < 130$ GeV. The expected yields for signal are calculated assuming $m_{\mathrm{H}} = 125$ GeV. [61]

events are seen than expected, whilst the observed numbers of 4μ and $2\mu2e$ agree very closely with those expected. The total number of expected events is within 1 standard deviation of that observed. Note that the uncertainties do not add in quadrature as the systematic uncertainties between samples are highly correlated.

The distributions $m_{4\ell}$ and σ_i from ATLAS 139 fb^{-1} data are compared with their predicted shapes from MC in Figure 7.6. Both predicted distributions are in agreement with those observed from data. As data and MC are in agreement, the model can be

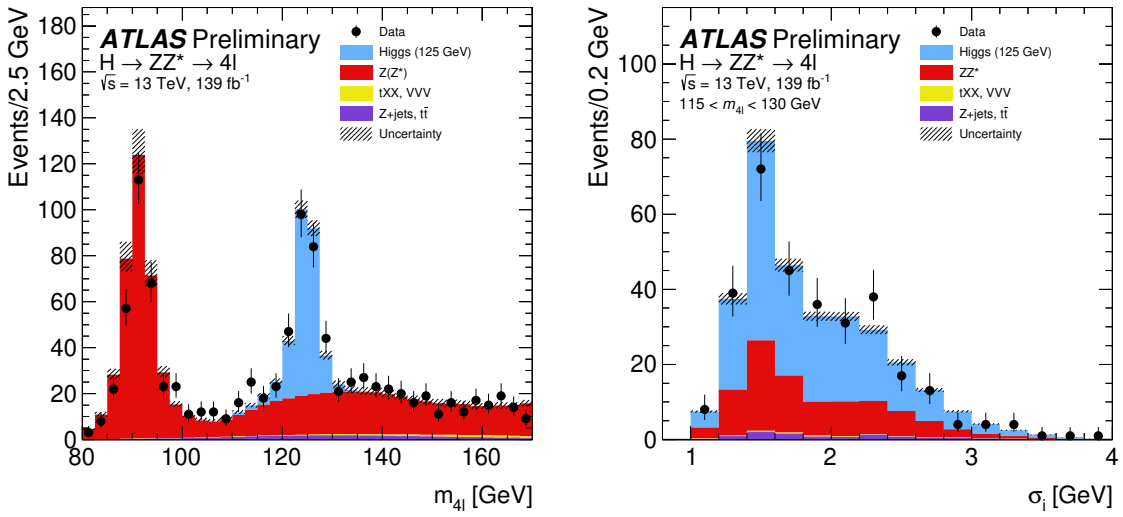


Fig. 7.6 Distributions of the $m_{4\ell}$ spectrum (left) and σ_i (right). The filled regions represent the expected distributions from MC simulation and the black points represent the observed distributions from data. In both cases, the MC and data are in agreement. [46]

fitted to the data.

7.2.1 PER model

Fitting the PER model to the data yields the result

$$m_{\mathrm{H}} = 124.92 \pm 0.19\ (\mathrm{stat.})\ {}^{+0.09}_{-0.06}\ (\mathrm{syst.})\ \mathrm{GeV},$$

which is in agreement with previous measurements by both ATLAS and CMS [6, 7, 45]. A projection of the fit across all 16 categories is shown in Figure 7.7. The results and profile likelihood scans from each channel are shown in Figure 7.8. The measurements from each channel are observed to be consistent with one another.

The observed statistical uncertainty of 187 MeV is 10 MeV higher than expected. The reason for this upward fluctuation can be seen in Table 7.2, which shows the relative normalisations. Slightly more background and less signal is seen than is expected, which will have the effect of increasing the uncertainty on m_{H}. From these numbers, the total signal and background normalisations are found to be 0.99 ± 0.08 and 1.09 ± 0.07 respectively. This small under-fluctuation is in agreement with the expected and observed numbers of events in Table 7.1, as the number of observed events is 0.98 of that expected. This also compares well with the observed global signal

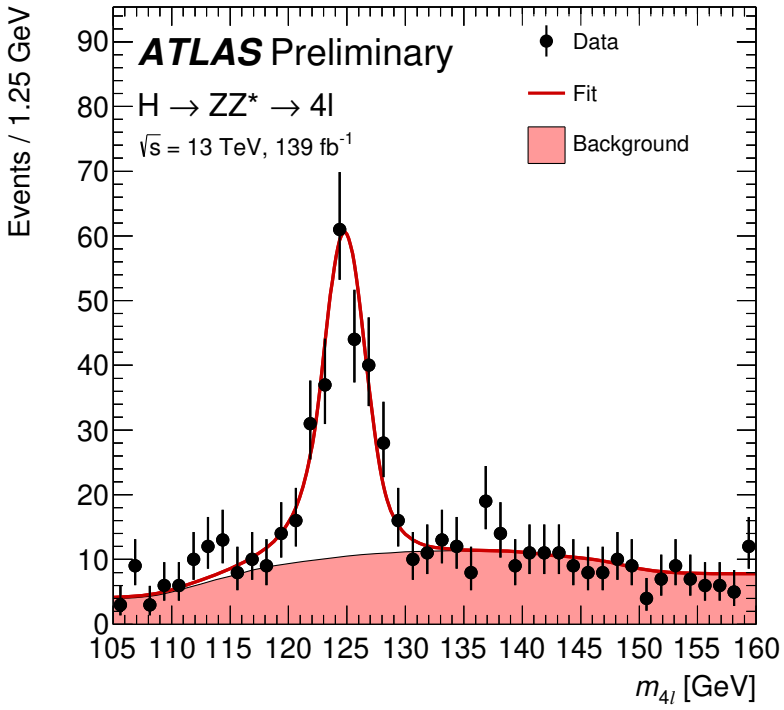


Fig. 7.7 Projection of the PER fitted across all 16 categories. The black points represent the ATLAS 139 fb^{-1} data, the red line, the fit, and the shaded region, the background model. [46]

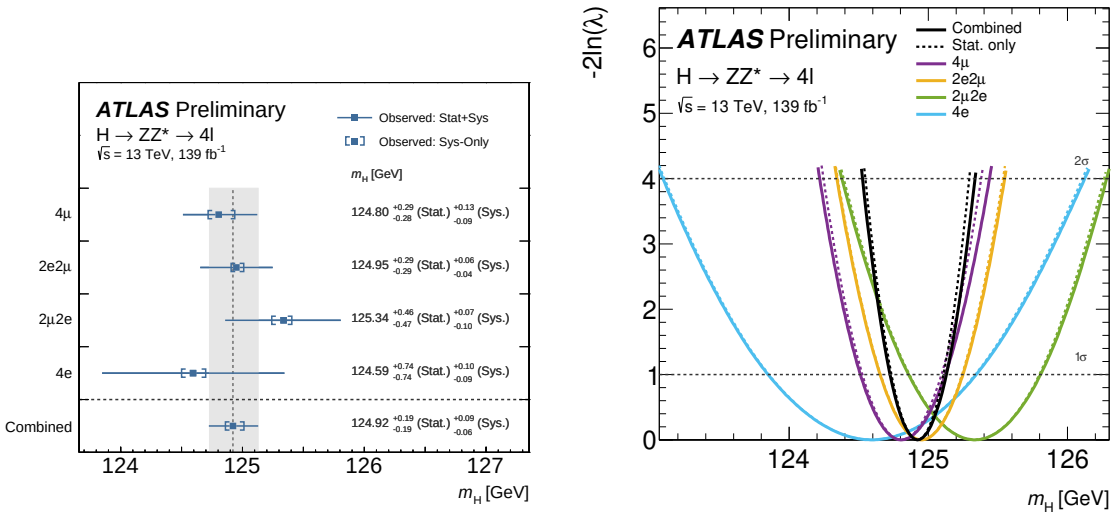


Fig. 7.8 Results from the PER model, per channel and combined. Shown are the measured values of m_{H} (left) and the likelihood scans (right) for each channel and combined. Each of the four per-channel fits is performed simultaneously across four BDT bins and the combined fit is performed simultaneously across all 16 categories. [46]

BDT bin	Relative normalisation
Signal	
1	$0.35^{+0.51}_{-0.47}$
2	$0.82^{+0.23}_{-0.21}$
3	0.86 ± 0.12
4	$1.30^{+0.15}_{-0.14}$
Background	
1	$1.08^{+0.12}_{-0.1}$
2	$1.06^{+0.14}_{-0.13}$
3	$1.14^{+0.14}_{-0.13}$
4	$1.09^{+0.23}_{-0.21}$

Table 7.2 Normalisations for signal and background for the PER model. A small excess of background and small under-fluctuation of signal are observed.

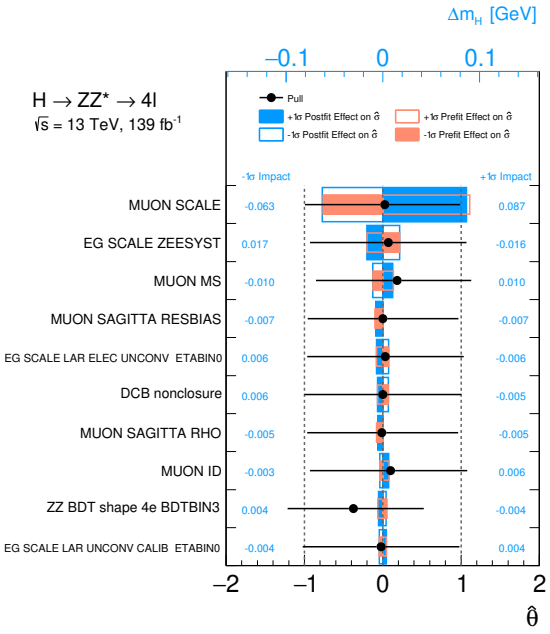


Fig. 7.9 Observed systematic ranking for the PER model.

strength in the ATLAS analysis of Higgs boson cross sections in $H \to ZZ^* \to 4\ell$ of 1.01 ± 0.11 [62]. The observed ranking of systematic uncertainties for the PER model is shown in Figure 7.9. This compares well with the expected ranking (Figure 7.2) as the two dominant systematics in both cases are the muon momentum scale and the electron energy scale.

7.2.2 Model without per-event resolution

Fitting the model without accounting for per-event resolution to the data yields the result

$$m_{\mathrm{H}} = 124.97 \pm 0.20 \text{ (stat.) } {}^{+0.09}_{-0.07} \text{ (syst.) GeV},$$

which is 51 MeV above that when accounting for per-event resolution. From Figure 7.5a, a 1σ difference between fitting the same data to the model with and without per-event resolution is 45 MeV, so the two values agree with a two-sided p-value of 26%. Taking the difference between the two uncertainties and using Figure 7.5b, they are found to be in agreement with a one-sided p-value of 17%. Figure 7.10 shows the distributions of uncertainties expected for the model when accounting for per-event resolution (black) and when not accounting for per-event resolution (blue). Also shown by the dashed vertical line are the observed uncertainties. The observed uncertainty for the PER model is found to be in agreement with that expected with a one-sided p-value of 29%.

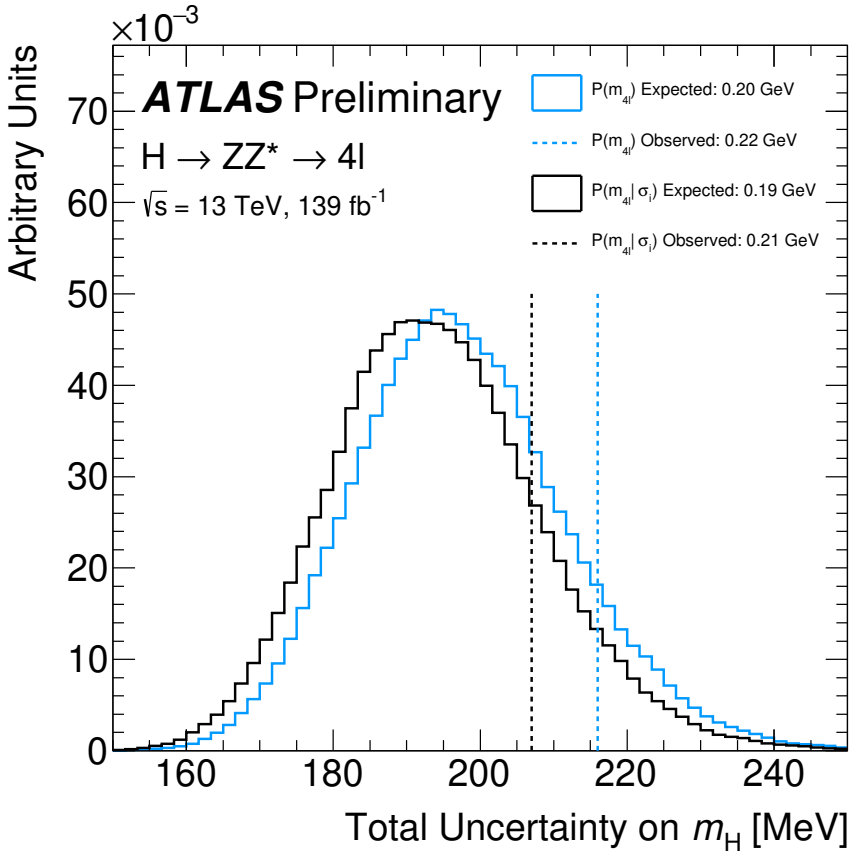


Fig. 7.10 Expected distribution of uncertainties on m_H when accounting for (black) and when not accounting for per-event resolution (blue). The solid lines indicate the distribution of uncertainties from the fits to bootstrapped MC sets and the dashed lines, the observed values of the uncertainties. [46]

Chapter 8

Conclusion

A new method to check the photon calibration procedure by transforming electrons from $Z \to ee$ decays into photons has been developed. From this it was found that the photon calibration procedure is performing as expected.

By transforming unconverted photons to single/double TRT converted and vice versa, the effect on the calibration of photon energy from misidentifying unconverted photons as single/double TRT converted and vice versa was investigated. It is found that misidentifying an unconverted photon as a single/double TRT converted leads to a 1% bias in photon energy calibration, independent of pseudorapidity. The reverse effect is found to be smaller, leading to a bias of -1% in the pseudorapidity region $0.8 \leq |\eta| < 1.37$. To determine the extent of the mis-calibration caused by this effect would require knowledge of the mis-identification rates.

To correct for the mismodelling of electromagnetic shower shapes in the ATLAS calorimeter, a new set of corrections to simulated electron showers has been determined. These corrections are used to improve the electron identification selections, which are a set of selections designed to remove "fake" electrons from data samples. A "fake" electron refers to either an electron from a hadronic decay or any other object misidentified as an electron. The selections were improved by updating the shift and width corrections to shower shape variables in simulated events. The corrections were determined from comparing simulated shower shapes to those observed in data for samples of $J/\psi \to ee$ and $Z \to ee$ decays. These corrections have been used in ATLAS publications.

A new system for applying shower-shape corrections by convolving the simulated distributions with a Gaussian has also been developed. This is found to over-correct in fewer cases than the method of shift and width corrections. Corrections using this method have not yet been used in publication.

An analytic description of the $m_{4\ell}$ spectrum from $H \to ZZ^* \to 4\ell$ events has been developed to measure the mass of the Higgs boson, using a Double-sided Crystal Ball model. A per-event description of the uncertainty is added to this model, using a Quantile-Regression Neural Network to determine the per-event uncertainty. When fitted to ATLAS 139 fb^{-1} data, the mass of the Higgs boson is determined to be

$$m_{\mathrm{H}} = 124.92 \pm 0.19 \text{ (stat.) } {}^{+0.09}_{-0.06} \text{ (syst.) GeV}.$$

This is in line with previous measurements by the ATLAS and CMS collaborations and represents the most precise measurement by ATLAS to date. It is also in agreement with the most recent CMS combination of $m_{\mathrm{H}} = 125.38 \pm 0.11$ (stat.) ± 0.09 (syst.) GeV to within two standard deviations [8].

The most recent ATLAS measurement of the Higgs boson cross section times branching ratio for the $H \to ZZ^*$ decay was 1.34 ± 0.12 pb, in agreement with the standard model prediction of 1.33 ± 0.08 pb [62]. For this measurement, $m_{\mathrm{H}} = 125$ GeV was assumed. As this mass is within the uncertainty of the measured mass in $H \to ZZ^* \to 4\ell$ and, from Figures 2.3 and 2.4, the Higgs boson branching ratio to ZZ^* and cross sections change very little between $m_{\mathrm{H}} = 124.92$ GeV and $m_{\mathrm{H}} = 125$ GeV, this measurement is compatible with the measured value of m_{H}.

References

[1] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[2] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[3] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". In: *Phys. Rev. Lett.* 13 (20 Nov. 1964), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.585.

[4] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.020. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.020.

[5] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (Sept. 2012), pp. 30–61. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.021. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.021.

[6] ATLAS Collaboration. "Measurement of the Higgs boson mass in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector". In: *Phys. Lett.* B784 (2018), pp. 345–366. DOI: 10.1016/j.physletb.2018.07.050. arXiv: 1806.00242 [hep-ex].

[7] CMS Collaboration. "Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV". In: *Journal of High Energy Physics* 2017.11 (Nov. 2017). ISSN: 1029-8479. DOI: 10.1007/jhep11(2017)047. arXiv: 1706.09936 [hep-ex]. URL: http://dx.doi.org/10.1007/JHEP11(2017)047.

[8] CMS Collaboration. "A measurement of the Higgs boson mass in the diphoton decay channel". In: *Physics Letters B* 805 (June 2020), p. 135425. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2020.135425. arXiv: 2002.06398 [hep-ex]. URL: http://dx.doi.org/10.1016/j.physletb.2020.135425.

[9] Mark Thomson. *Modern Particle Physics*. Cambridge University Press, 2013.

[10] Francis Halzen and Alan D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley and Sons, 1984.

[11] Konstantinidis, N. et al. *Lecture notes for the 2017 HEP School for Experimental High Energy Physics Students*. Tech. rep. Science and Technology Facilities Council, 2017. URL: http://purl.org/net/epubs/manifestation/34993301/RAL-TR-2017-009.pdf.

[12] Abdelhak Djouadi. "The anatomy of electroweak symmetry breaking". In: *Physics Reports* 457.1-4 (Feb. 2008), pp. 1–216. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2007.10.004. URL: http://dx.doi.org/10.1016/j.physrep.2007.10.004.

[13] Particle Data Group. "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[14] M. et al. Aker. "Improved Upper Limit on the Neutrino Mass from a Direct Kinematic Method by KATRIN". In: *Physical Review Letters* 123.22 (Nov. 2019). ISSN: 1079-7114. DOI: 10.1103/physrevlett.123.221802. URL: http://dx.doi.org/10.1103/PhysRevLett.123.221802.

[15] LHC Higgs Cross Section Working Group. *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*. CERN Yellow Reports: Monographs. Oct. 2016. DOI: 10.23731/CYRM-2017-002. URL: https://cds.cern.ch/record/2227475.

[16] Lyndon Evans and Philip Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08001–S08001. DOI: 10.1088/1748-0221/3/08/s08001. URL: https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08001.

[17] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3.S08003 (2008). DOI: 2008JINST3S08003.

[18] Karolos Potamianos. *The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider*. Tech. rep. ATL-PHYS-PROC-2016-104. 15 pages, EPS-HEP 2015 Proceedings. Geneva: CERN, Aug. 2016. URL: https://cds.cern.ch/record/2209070.

[19] Richard Wigmans. "Calorimetry". In: *Scientifica Acta 2* 1 (2008), pp. 18–55. URL: http://siba.unipv.it/fisica/ScientificaActa/volume_2_1/Wigmans.pdf.

[20] J H Hubbell. "Review of photon interaction cross section data in the medical and biological context". In: *Physics in Medicine and Biology* 44.1 (Jan. 1999), R1–R22. DOI: 10.1088/0031-9155/44/1/001. URL: https://doi.org/10.1088%2F0031-9155%2F44%2F1%2F001.

[21] Richard Wigmans and Mehmet T Zeyrek. "On differences between calorimetric detection of electrons and photons". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 485.3 (2002), pp. 385–398. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(01)02141-6. URL: http://www.sciencedirect.com/science/article/pii/S0168900201021416.

[22] W. Bonivento. "Overview of the ATLAS electromagnetic calorimeter". In: *Nuclear Physics B - Proceedings Supplements* 78.1 (1999). Advanced Technology and Particle Physics, pp. 176–181. ISSN: 0920-5632. DOI: https://doi.org/10.1016/S0920-5632(99)00542-3. URL: http://www.sciencedirect.com/science/article/pii/S0920563299005423.

[23] ATLAS Collaboration. "Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data". In: *Journal of Instrumentation* 14.03 (Mar. 2019), P03017–P03017. DOI: 10.1088/1748-0221/14/03/p03017. URL: https://doi.org/10.1088%2F1748-0221%2F14%2F03%2Fp03017.

[24] Henric Wilkens (on behalf of the ATLAS LArg Collaboration). "The ATLAS Liquid Argon calorimeter: An overview". In: *J. Phys.* Conf. Ser. 160 012043 (2009).

[25] ATLAS Collaboration. "Performance of the ATLAS Trigger System in 2015". In: *The European Physical Journal C* 77.5 (May 2017). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-4852-3. arXiv: 1611.09661 [hep-ex]. URL: http://dx.doi.org/10.1140/epjc/s10052-017-4852-3.

[26] Aranzazu Ruiz-Martinez and ATLAS Collaboration. *The Run-2 ATLAS Trigger System*. Tech. rep. ATL-DAQ-PROC-2016-003. Geneva: CERN, Feb. 2016. DOI: 10.1088/1742-6596/762/1/012003. URL: https://cds.cern.ch/record/2133909.

[27] ATLAS Collaboration. "Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data". In: *Journal of Instrumentation* 14.12 (Dec. 2019), P12006–P12006. ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/12/p12006. arXiv: 1908.00005 [hep-ex]. URL: http://dx.doi.org/10.1088/1748-0221/14/12/P12006.

[28] Zachary Marshall and. "Simulation of Pile-up in the ATLAS Experiment". In: *Journal of Physics: Conference Series* 513.2 (June 2014), p. 022024. DOI: 10.1088/1742-6596/513/2/022024. URL: https://doi.org/10.1088%2F1742-6596%2F513%2F2%2F022024.

[29] W Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. Tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003. Geneva: CERN, Apr. 2008. URL: https://cds.cern.ch/record/1099735.

[30] ATLAS Collaboration. "Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2". In: *The European Physical Journal C* 77.10 (Oct. 2017). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-5225-7. arXiv: 1704.07983 [hep-ex]. URL: http://dx.doi.org/10.1140/epjc/s10052-017-5225-7.

[31] ATLAS Collaboration. "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV". In: *The European Physical Journal C* 79.8 (Aug. 2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7140-6. URL: http://dx.doi.org/10.1140/epjc/s10052-019-7140-6.

[32] ATLAS Collaboration. "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data". In: *The European Physical Journal C* 74.10 (Oct. 2014). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-014-3071-4. URL: http://dx.doi.org/10.1140/epjc/s10052-014-3071-4.

[33] Byron Roe and Haijun Yang. "Boosted decision trees, a powerful event classifier". In: (May 2006). DOI: 10.1142/9781860948985_0029.

[34] M. Aaboud et al. "Measurements of Higgs boson properties in the diphoton decay channel with 36fb1 of pp collision data at s=13TeV with the ATLAS detector". In: *Physical Review D* 98.5 (Sept. 2018). ISSN: 2470-0029. DOI: 10.1103/physrevd.98.052005. URL: http://dx.doi.org/10.1103/PhysRevD.98.052005.

[35] A. Hoecker et al. *TMVA - Toolkit for Multivariate Data Analysis*. 2007. arXiv: physics/0703039 [physics.data-an].

[36] Kyle Cranmer. "Kernel estimation in high-energy physics". In: *Computer Physics Communications* 136.3 (May 2001), pp. 198–207. ISSN: 0010-4655. DOI: 10.1016/s0010-4655(00)00243-5. URL: http://dx.doi.org/10.1016/S0010-4655(00)00243-5.

[37] ATLAS Collaboration. "Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector". In: *Phys. Rev. D* 90 (5 Sept. 2014), p. 052004. DOI: 10.1103/PhysRevD.90.052004. URL: https://link.aps.org/doi/10.1103/PhysRevD.90.052004.

[38] ATLAS Collaboration. "Measurement of inclusive and differential cross sections in the $H \to ZZ^* \to 4\ell$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 10 (2017), p. 132. DOI: 10.1007/JHEP10(2017)132. arXiv: 1708.02810 [hep-ex].

[39] G. Aad et al. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV". In: *The European Physical Journal C* 76.5 (May 2016). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex]. URL: http://dx.doi.org/10.1140/epjc/s10052-016-4120-y.

[40] ATLAS Collaboration. "Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector". In: *Phys. Rev. D* 91 (1 Jan. 2015), p. 012006. DOI: 10.1103/PhysRevD.91.012006. URL: https://link.aps.org/doi/10.1103/PhysRevD.91.012006.

[41] Graham John Cree. "Direct measurement of the Higgs boson mass, natural width, and cross section times branching ratio to four leptons using a per-event lineshape in the Higgs to ZZ to four lepton decay channel with the ATLAS detector". Presented 19 Apr 2017. PhD thesis. Carleton University, Apr. 2017. URL: https://cds.cern.ch/record/2286849.

[42] Valerio Ippolito. "Measurement of the properties of the new particle observed within the search for the Higgs boson in the $H \to ZZ^{(*)} \to 4\ell$ channel at ATLAS." Presented 18 Feb 2014. PhD thesis. University of Rome, Dec. 2013. URL: http://cds.cern.ch/record/1966687.

[43] LHC Higgs Cross Section Working Group et al. *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*. 2011. arXiv: 1101.0593 [hep-ph].

[44] CMS Collaboration. "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV". In: *The European Physical Journal C* 75.5 (May 2015). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-015-3351-7. arXiv: 1412.8662 [hep-ex]. URL: http://dx.doi.org/10.1140/epjc/s10052-015-3351-7.

[45] The ATLAS and CMS Collaborations. "Combined Measurement of the Higgs Boson Mass in *pp* Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments". In: *Physical Review Letters* 114.19 (May 2015). ISSN: 1079-7114. DOI: 10.1103/physrevlett.114.191803. URL: http://dx.doi.org/10.1103/PhysRevLett.114.191803.

[46] ATLAS Collaboration. *Measurement of the Higgs boson mass in the $H \to ZZ^* \to 4\ell$ decay channel with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector at the LHC*. Tech. rep. ATLAS-CONF-2020-005. Geneva: CERN, Apr. 2020. URL: http://cds.cern.ch/record/2714883.

[47] Olaf Behnke et al. *Data Analysis in High Energy Physics: A Practical Guide to Statistical Methods*. 1st. Wiley-VCH, 2013. ISBN: 3527410589.

[48] Christoph Langenbruch. *Parameter uncertainties in weighted unbinned maximum likelihood fits*. 2019. arXiv: 1911.01303 [physics.data-an].

[49] Wouter Verkerke and David Kirkby. *The RooFit toolkit for data modeling*. 2003. arXiv: physics/0306116 [physics.data-an].

[50] M. Oreglia. "A Study of the Reactions $\psi' \to \gamma\gamma\psi$". PhD thesis. SLAC, 1980. URL: http://www-public.slac.stanford.edu/sciDoc/docMeta.aspx?slacPubNumber=slac-r-236.html.

[51] J. Gaiser. "Charmonium spectroscopy from radiative decays of the J/ψ and ψ'". In: *Ph.D. Thesis, SLAC-R-255* (1982). Appendix F., SLAC.

[52] ATLAS Collaboration. *Search for Scalar Diphoton Resonances in the Mass Range $65 - 600$ GeV with the ATLAS Detector in pp Collision Data at $\sqrt{s} = 8$ TeV*. Oct. 2014. DOI: 10.1103/physrevlett.113.171801. arXiv: 1407.6583 [hep-ex]. URL: http://dx.doi.org/10.1103/PhysRevLett.113.171801.

[53] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *The European Physical Journal C* 71.2 (Feb. 2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1554-0. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1554-0.

[54] David B. Fogel Zbigniew Michalewicz. *How to Solve It: Modern Heuristics*. 2nd ed. Springer Science Business Media, Mar. 2013. ISBN: 978-3-662-07807-5.

[55] Luc Demortier and Louis Lyons. *Everything you always wanted to know about pulls*. Tech. rep. CDF/ANAL/PUBLIC/5776. The Rockefeller University,University of Oxford, Apr. 2002. URL: http://physics.rockefeller.edu/luc/technical_reports/cdf5776_pulls.pdf.

[56] Alex J. Cannon. "Quantile regression neural networks: Implementation in R and application to precipitation downscaling". In: *Computers Geosciences* 37.9 (2011), pp. 1277–1284. DOI: 10.1016/j.cageo.2010.07.005. URL: http://www.sciencedirect.com/science/article/pii/S009830041000292X.

[57] Martin Abadi et al. "TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems". In: (2015). Software available from tensorflow.org. URL: https://www.tensorflow.org/.

[58] François Chollet et al. "Keras". In: (2015).

[59] M. Baak et al. "Interpolation between multi-dimensional histograms using a new non-linear moment morphing method". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 771 (2015), pp. 39–48. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2014.10.033. URL: http://www.sciencedirect.com/science/article/pii/S0168900214011814.

[60] ATLAS Collaboration. "Combined search for the Standard Model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Physical Review D* 86.3 (Aug. 2012). ISSN: 1550-2368. DOI: 10.1103/physrevd.86.032003. arXiv: 1207.0319 [hep-ex]. URL: http://dx.doi.org/10.1103/PhysRevD.86.032003.

[61] ATLAS Collaboration. "Measurements of the Higgs boson inclusive and differential fiducial cross sections in the 4ℓ decay channel at $\sqrt{s} = 13$ TeV". In: (2020). arXiv: 2004.03969 [hep-ex].

[62] ATLAS Collaboration. *Higgs boson production cross-section measurements and their EFT interpretation in the 4ℓ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector*. 2020. arXiv: 2004.03447 [hep-ex].